IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:	Confirmation No.: 5386
Gerard *et al.*	Art Unit: 1652
Appl. No. 09/064,057	Examiner: Nashed, Nashaat T.
Filing date: April 22, 1998	Atty. Docket: 0942.4330002/RWE/HCC
For: **Recombinant Methods for Making Reverse Transcriptases and Mutants Thereof**	

Brief on Appeal Under 37 C.F.R. § 41.37

Mail Stop Appeal Brief - Patents

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

A Notice of Appeal from the final rejection of claims 26, 33, 117-125, 137-147 and 149-161 was filed on February 22, 2005. Appellant hereby files this Appeal Brief, together with the required brief filing fee.

It is not believed that extensions of time are required beyond those that may otherwise be provided for in documents accompanying this paper. However, if additional extensions of time are necessary to prevent abandonment of this application, then such extensions of time are hereby petitioned under 37 C.F.R. § 1.136(a), and any fees required therefor (including fees for net addition of claims) are hereby authorized to be charged to our Deposit Account No. 19-0036.

Table of Contents

I. Real Parties In Interest

The real party in interest in this appeal is Invitrogen Corporation.

II. Related Appeals and Interferences

No other prior or pending appeals, interferences or judicial proceedings are known to the Appellant, the Appellant's legal representative, or assignee which may be related to, or directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

III. Status of Claims

Claims 26, 33, 117, 122-125, 137-147 and 149-161 are pending in the application.

Claims 1-25, 27-32, 34-116, 118-121, 126-136 and 148 have been canceled.

Claims 26, 33, 117-125, 137-147 and 149-161 are rejected.

IV. Status of Amendments

Except for the cancellation of claims 118-121 (*see* the Supplemental Amendment and Reply Under 37 C.F.R § 1.116 filed September 2, 2004), all amendments have been entered.[1]

---

[1] Appellant directs the Board's attention to the fact that the Examiner previously acknowledged the cancellation of claims 118-121 in the Advisory Action dated May 24, 2004, page 2. *See* Advisory Action dated May 24, 2004, page 2. However, in the Final Office Action dated October 25, 2004, the Examiner states that 26, 33, 117-125, 137-147

V. Summary of Claimed Subject Matter

The present invention relates generally to methods of producing an Avian Myeloblastosis Virus (AMV) reverse transcriptase having a high RNA dependent DNA polymerase specific activity. In particular, the methods require the co-expression of an α subunit of AMV reverse transcriptase and an β subunit of AMV reverse transcriptase in a eukaryotic host cell, wherein the eukaryotic host cell comprises one or more nucleic acid sequences encoding an AMV reverse transcriptase α subunit and one or more nucleic acid sequences encoding an AMV reverse transcriptase β subunit, and the isolation or purification of the expressed AMV reverse transcriptase, thereby obtaining an RNA dependent DNA polymerase specific activity of at least about 30,000 units per milligram. *See* specification at page 26, lines 5-25, and at page 28, line 29 through page 29, line 27. Thus, the invention is to methods involving the co-expression of the α subunit of AMV reverse transcriptase and β subunit of AMV reverse transcriptase, *together*, in the same eukaryotic host cell to produce an AMV reverse transcriptase having an RNA dependent DNA polymerase specific activity of at least about 30,000 units per milligram

Claim 26 is the sole independent claim pending in the present application. Claim 26 defines a method for the production of AMV reverse transcriptase having an RNA dependent DNA polymerase specific activity of at least about 30,000 units per milligram. The method comprises obtaining a eukaryotic host cell comprising one or more nucleic acid sequences encoding an AMV reverse transcriptase α subunit and one or more nucleic acid sequences encoding an AMV reverse transcriptase β subunit; culturing said

---

and 149-161 are pending and under consideration. See Final Office Action dated October 25, 2004, page 2.

host cell in conditions appropriate to produce said AMV reverse transcriptase; and isolating or purifying said AMV reverse transcriptase such that it has a specific activity of at least about 30,000 units per milligram.

Support for claim 26 can be found throughout the specification, for example, at page 12, lines 15-21, at page 26, line 5 through page 27, line 5, and at page 28, line 29 through page 29, line 27.

VI. Grounds of Rejection to be Reviewed on Appeal

Claims 26, 33, 117-125, 137-147 and 149-161 stand rejected under 35 U.S.C. § 103, as being unpatentable over Soltis *et al.*, *Proc. Natl. Acad. Sci. USA 85:*3372-3376 (1988) ("Soltis")[2], in view of the state of the art at the time the application was filed, as allegedly exemplified by Kawa *et al.*, *Protein Expr. Purif. 4:*298-303 (1993) ("Kawa")[3], Barr *et al.*, *Bio/Technology 5:*486-489 (1987) ("Barr")[4], and Ford *et al.*, *Protein Expr. Purif. 2:*95-107 (1991) ("Ford")[5]. *See* Final Office Action dated October 25, 2004 ("Final Office Action"), page 2. Appellant has traversed this rejection.

---

[2] Soltis is submitted herein as Exhibit 1, in compliance with 37 C.F.R. § 41.37(c)(ix).

[3] Kawa is submitted herein as Exhibit 2, in compliance with 37 C.F.R. § 41.37(c)(ix).

[4] Barr is submitted herein as Exhibit 3, in compliance with 37 C.F.R. § 41.37(c)(ix).

[5] Ford is submitted herein as Exhibit 4, in compliance with 37 C.F.R. § 41.37(c)(ix).

VII. Argument

A. Legal Standard for Obviousness

In proceedings before the Patent and Trademark Office, the Examiner bears the burden of establishing a prima facie case of obviousness based upon the prior art. *See* In re Piasecki, 745 f.2d 1468, 1471-73 (Fed. Cir. 1984). To meet this burden, the Examiner must satisfy three requirements. First, all of the claim limitations must be taught or suggested by the prior art. *See* In re Royka, 490 F.2d 981, 984-85 (CCPA 1974); *see also* In re Glaug , 283 F.3d 1335, 1341-42 (Fed. Cir. 2002) (finding a claim not obvious because the prior art did not teach "spaced zones of adhesive" as recited in the claim); In re Rijckaert, 9 F.3d 1531, 1533 (Fed. Cir. 1993) (finding a claim not obvious because the prior art did not teach all claim limitations). Second, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. *See* In re Rouffet, 149 F.3d 1350, 1357 (Fed. Cir. 1998). Third, there must be a reasonable expectation of success. *See* In re Merck & Co., Inc., 800 F.2d 1091 (Fed. Cir. 1986). The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, not in Applicants' disclosure. *See* In re Vaeck, 947 F.2d 488 (Fed. Cir. 1991).

Evidence of a suggestion, teaching, or motivation to combine prior art references may flow, *inter alia*, from the references themselves, the knowledge of one of ordinary skill in the art, or from the nature of the problem to be solved. *See* In re Dembiczak, 175 F.3d 994, 999 (Fed. Cir. 1999). Although a reference need not expressly teach that the

disclosure contained therein should be combined with another, *see* Motorola, Inc. v. Interdigital Tech. Corp., 121 F.3d 1461, 1472 (Fed. Cir. 1997), the showing of combinability, in whatever form, must nevertheless be "clear and particular." Dembiczak, 175 F.3d at 999. "Broad conclusory statements regarding the teaching of multiple references, standing alone, are not 'evidence.'" *Id.* at 999; *see also* In re Kotzab, 217 F.3d 1365, 1371 (Fed. Cir. 2000) ("particular findings must be made as to the reason the skilled artisan, with no knowledge of the claimed invention, would have selected these components for combination in the manner claimed.")

B. The Subject Matter of Claims 26, 33, 117-125, 137-147 and 149-161 is Not Obvious Over the Cited References

In the remarks that follow, Appellant focuses solely on independent claim 26, as any claim that depends from a patentable independent claim is also patentable by virtue of its dependency. Claim 26 is directed to a method of producing an Avian Myeloblastosis Virus (AMV) reverse transcriptase by co-expressing nucleic acid sequences encoding the α and β subunits of AMV reverse transcriptase, together, in a eukaryotic cell, and isolating or purifying the expressed reverse transcriptase, thereby obtaining an AMV reverse transcriptase having a specific activity of at least about 30,000 units per milligram. Thus, claim 26 recites the following elements: (1) expressing the α subunit, and (2) expressing the β subunit, (3) in the same eukaryotic cell comprising one or more nucleic acids sequences encoding an α subunit and one or more nucleic acids sequences encoding a β subunit, (4) isolating or purifying the AMV reverse transcriptase, and (5) obtaining a specific activity of $\geq$ 30,000 units per milligram.

Soltis discloses the *independent* expression of the α and β subunits of avian retrovirus reverse transcriptase in *E. coli*. More specifically, Soltis discloses the individual expression of the Avian Sarcoma and Leukosis Virus (ASLV) α subunit alone in an *E. coli* host cell, the individual expression of the ASLV β subunit alone in an *E. coli* host cell, and the separate purification of each distinct subunit. The resulting subunits were shown to have a specific activity that was 10,000 to 100,000 fold *lower* than the specific activity of AMV reverse transcriptase purified directly from AMV particles. *See* page 3373, right column. Thus, Soltis discloses the independent expression of the α and β subunits of an avian retrovirus reverse transcriptase in a prokaryotic host cell with the resultant reverse transcriptase having substantially lower specific activity than that of an AMV reverse transcriptase purified directly from AMV particles. As such, Soltis does not teach or suggest the co-expression of the α and β subunits of AMV reverse transcriptase, *together*, in a *eukaryotic* cell, and the isolation or purification of the expressed AMV reverse transcriptase, thereby obtaining an AMV reverse transcriptase has a high a specific activity, *e.g.*, a specific activity of at least about 30,000 units per milligram.

Kawa discloses the expression and purification of Human Immunodeficiency virus-1 (HIV-1) reverse transcriptase using a baculovirus expression system (BEVS). Kawa discloses the expression of the HIV-1 *pol* gene in Sf9 cells. The methods of Kawa yielded a homodimer of the p66 subunit of HIV-1 reverse transcriptase having DNA polymerase activity but did not produce the mature p66/p55 heterodimer form of the enzyme. *See* abstract. Kawa also discloses that the baculovirus-produced HIV-1 reverse transcriptase had a *lower* specific activity than that derived from *E.* coli. *See* page 302,

left column. Kawa does not disclose or suggest the co-expression of the α subunit and β subunit of AMV reverse transcriptase in a single eukaryotic cell having one or more nucleic acids sequences encoding an α subunit and one or more nucleic acids sequences encoding a β subunit. Further, Kawa is totally void of any reference to AMV reverse transcriptase. HIV-1 reverse transcriptase is a distinct chemical entity from AMV reverse transcriptase and has a different chemical structure and amino acid sequence than AMV reverse transcriptase. A person of ordinary skill in the art would therefore not expect that results observed in the production of HIV-1 reverse transcriptase in a BEVS expression system would have any relevance to the production of an AMV reverse transcriptase having high specific activity. Thus, there is no sound basis on which to assert that results provided in Kawa would be representative of the state of the art for the production of AMV reverse transcriptase having high specific activity. As such, Kawa does not teach or suggest the present invention and certainly does not cure the deficiencies of Soltis. If fact, assuming *arguendo*, if one of ordinary skill in the art were to combine the teachings of Soltis and Kawa, they would be led away from a reasonable expectation of success in obtaining an AMV reverse transcriptase with a high specific activity.

Barr discloses the expression of active HIV reverse transcriptase in *S. cerevisiae*. Specifically, Barr discloses the expression of the reverse transcriptase domain of the HIV *pol* gene in yeast cells. *See* abstract. While the methods of Barr yielded both the p66 subunit and the p51 subunit of the mature form of the enzyme, Barr does not teach or suggest the co-expression of the α subunit and β subunit of AMV reverse transcriptase in the same eukaryotic cell which has one or more nucleic acids sequences encoding an α

subunit and one or more nucleic acids sequences encoding a β subunit. In addition, Barr does not teach or suggest that an AMV reverse transcriptase could be obtained with a high specific activity. In fact, Barr is totally void of any reference to AMV reverse transcriptase. Barr's disclosure is limited to HIV-1 reverse transcriptase, which is a distinct chemical entity from AMV reverse transcriptase and has a different chemical structure and amino acid sequence than AMV reverse transcriptase. Moreover, Barr expressed a single gene whose product is further processed to give a heterodimer consisting of equimolar amounts of the two species designated p66 and p51. In the present method, two distinct genes are used to express the two distinct proteins that make up the subunits of AMV reverse transcriptase. Thus, a person of ordinary skill in the art would not expect that the results observed in the production of HIV-1 reverse transcriptase in *S. cerevisiae* would have any relevance to the production of AMV reverse transcriptase. Therefore, there is no sound basis on which to assert that the disclosure provided in Barr would be representative of the state of the art for the production of an AMV reverse transcriptase having a high specific activity.

Ford discloses the use of fusion tails for the recovery and purification of recombinant proteins. Ford provides a review of several fusion tail systems which include genetically engineering a target protein to contain a C- or N- terminal polypeptide, which provides the biochemical basis for specificity in recovery and purification of the target protein. Ford, like Kawa and Barr, is completely void of any reference to AMV reverse transcriptase. Moreover, Ford does not make any reference to, or suggestion that, AMV reverse transcriptase would be a suitable target protein for use in a fusion tail system. Thus, Ford does not provide guidance in the co-expression of

the α and β subunits of AMV reverse transcriptase, *together*, in a *eukaryotic* cell, and the isolation or purification of the expressed AMV reverse transcriptase, thereby giving an AMV reverse transcriptase having a specific activity of at least about 30,000 units per milligram as required by pending claim 26 of the present application.

As explained *supra*, the art cited by the Examiner, when taken alone or in combination, fails to teach or suggest all the of the claim elements of Claim 26. An obviousness rejection, among other things, requires that *all* of the claim elements be taught or suggested by the cited references. *See* In re Royka at 984-85. Since all elements of claim 26 are not taught or suggested by the references relied on in the rejection, claim 26, and claims 33, 117, 122-125, 137-147 and 149-161 that depend from claim 26, are not obvious over the references cited by the Examiner.

In addition, the Examiner has not met his burden of providing an objective suggestion or motivation to combine or modify the cited references to obtain the claimed invention. *See* In re Fine, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988). Instead, the Examiner relies on the conclusory statement that "AMV reverse transcriptase has been used extensively in biotechnology in the preparation of cDNA libraries. Thus, one of ordinary skill in the art would had motivation to develop a method to produce AMV reverse transcriptase by a recombinant method." Final Office Action, page 3. In an attempt to buttress this conclusory statement, the Examiner incorrectly asserts that Soltis provides motivation for the co-expression of the α and β subunits in the same host cell because Soltis discloses that the heterodimer has higher thermostability than the homodimer. *See* Final Office Action, page 3. However, a close reading of Soltis shows that the

bacterially expressed α homodimer protein could be distinguished from both the AMV reverse transcriptase *and* the β homodimer by its increased sensitivity to heat inactivation. These results, as stated in Soltis at page 3376, left column, merely support previous observations regarding the α subunit purified from ASLV virions and provide no motivation produce a heterodimeric AMV reverse transcriptase. In addition, the Appellant asserts that a person of ordinary skill in the art would not have been motivated to combine disclosures of Kawa and Barr, which are directed to HIV reverse transcriptase, with Soltis to cure the deficiencies of the Soltis disclosure. As discussed *supra*, HIV reverse transcriptase is a distinct chemical entity from that of AMV reverse transcriptase and has a different chemical structure and amino acid sequence than AMV reverse transcriptase. For at least these reasons, the Examiner has failed to meet his burden of establishing "some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings." *See* In re Rouffet, 149 F.3d at 1357. Accordingly, the rejection of claims 26, and claims 33, 117, 122-125, 137-147 and 149-161 that depend from claim 26, under 35 U.S.C. § 103 is improper and should be reversed.

Further, a person of ordinary skill in the art would not have had a reasonable expectation of success in combining the references cited by the Examiner to yield the claimed invention. *See* In re Merck & Co., Inc., 800 F.2d 1091. Soltis discloses the independent expression of the α subunit and the β subunit of ASLV reverse transcriptase in a prokaryotic host cell. Kawa and Barr disclose the production of an active form of HIV reverse transcriptase in insect cells and yeast cells by expressing the full-length *pol*

gene of HIV reverse transcriptase. A person of ordinary skill in the art would not have reasonably expected that the independent expression of the α subunit and the β subunit of ASLV reverse transcriptase in a prokaryotic host cell of Soltis could have been combined with the expression of the full-length *pol* gene of HIV reverse transcriptase in a eukaryotic host cell of Kawa and Barr to successfully yield a method of co-expression of an α subunit of AMV reverse transcriptase and an β subunit of AMV reverse transcriptase, together, in a eukaryotic host cell having one or more nucleic acid sequences encoding AMV reverse transcriptase α subunit and one or more nucleic acid sequences encoding an AMV reverse transcriptase β subunit. In addition, Soltis discloses that the specific activity of the recombinant ASLV reverse transcriptase was substantially lower than that of purified AMV reverse transcriptase. Likewise, Kawa discloses that the specific activity of the eukaryotically derived recombinant HIV-1 reverse transcriptase was lower than the prokaryotically derived recombinant form. Barr fails to disclose any comparison of the specific activity of the recombinant HIV reverse transcriptase with either the naturally occurring form or a prokaryotically derived recombinant form of HIV reverse transcriptase. Thus, a person of ordinary skill in the art would have had no reasonable expectation of success, based on the references cited by the Examiner, in obtaining an AMV reverse transcriptase with high specific activity. Since there is no suggestion or motivation to combine or modify the references relied on by the Examiner, and since there would have been no reasonable expectation of success in combining or modifying the cited references to yield the claimed method, the subject matter of claims 26, and claims 33, 117, 122-125, 137-147 and 149-161 that depend from claim 26, cannot be said to be obvious over the cited references.

C. Conclusion

In view of the foregoing discussion, Appellant respectfully submits that the subject matter defined by claims 26, 33, 117, 122-125, 137-147 and 149-161 is patentable over the cited art and that the Examiner has failed to meet his burden of establishing a *prima facie* case of obviousness. Accordingly, Appellant respectfully requests that the Board reverse the Examiner's final rejection of these claims under 35 U.S.C. § 103 and remand this application for issue.

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

Robert W. Esmond
Attorney for Appellant
Registration No. 32,893

Date: April 21, 2005

1100 New York Avenue, N.W.
Washington, D.C. 20005-3934
(202) 371-2600

385892_2.DOC

VIII. Claims Appendix

26. A method of producing an Avian Myeloblastosis Virus (AMV) reverse transcriptase having an RNA dependent DNA polymerase specific activity of at least about 30,000 units per milligram, said method comprising

(a) obtaining a eukaryotic host cell comprising one or more nucleic acid sequences encoding an AMV reverse transcriptase α subunit and one or more nucleic acid sequences encoding an AMV reverse transcriptase β subunit; and

(b) culturing said host cell under conditions sufficient to produce said AMV reverse transcriptase; and

(c) isolating or purifying said reverse transcriptase thereby obtaining an AMV reverse transcriptase having an RNA-dependent DNA polymerase specific activity of at least about 30,000 units per milligram.

33. The method of claim 26, wherein subunits of said AMV reverse transcriptase are expressed in said host cell to form said AMV reverse transcriptase.

117. The method of claim 26, wherein said nucleic acid sequences are contained in one or more vectors.

122. The method of claim 26, wherein said AMV reverse transcriptase comprises two β subunits.

123. The method of claim 26, wherein said AMV reverse transcriptase comprises an α and a β subunit.

124. The method of claim 117, wherein said nucleic acid sequences are contained on the same vector.

125. The method of claim 117, wherein said nucleic acid sequences are contained on different vectors.

137. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity from about 80,000 units per milligram to about 150,000 units per milligram.

138. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 35,000 units per milligram.

139. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 40,000 units per milligram.

140. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 45,000 units per milligram.

141. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 50,000 units per milligram.

142. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 55,000 units per milligram.

143. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 60,000 units per milligram.

144. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 65,000 units per milligram.

145. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 70,000 units per milligram.

146. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 75,000 units per milligram.

147. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of at least about 80,000 units per milligram.

149. The method of claim 26, wherein said host cell is a cultured insect cell.

150. The method of claim 26, wherein said host cell is an insect larva cell.

151. The method of claim 26, wherein said host cell is a yeast cell.

152. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 35,000 units per milligram.

153. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 40,000 units per milligram.

154. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 45,000 units per milligram.

155. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 50,000 units per milligram.

156. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 55,000 units per milligram.

157. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 60,000 units per milligram.

158. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 65,000 units per milligram.

159. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 70,000 units per milligram.

160. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 75,000 units per milligram.

161. The method of claim 26, wherein said AMV reverse transcriptase has an RNA-dependent DNA polymerase specific activity of about 80,000 units per milligram.

IX. Evidence Appendix

Exhibit 1 Soltis, D. et al., "The α and β Chains of Avian Retrovirus reverse Transcriptase Independently Expressed In *Escherichia coli*: Characterization of Enzymatic Activities," Proc. Natl. Acad. Sci. USA 85:3372-3376 (1988)

Exhibit 2 Kawa, S., et al., "Expression and Purification of the HIV-1 Reverse Transcriptase Using the Baculovirus Expression Vector System," Protein Expr. Purif. 4:298-303 (1993).

Exhibit 3 Barr, P., et al., "Expression of Active Human Immunodeficiency Virus Reverse Transcriptase in *Saccharomyces Cerevisiae*," Bio/Technology 5:486-489 (1987).

Exhibit 4 Ford, C., et al., "Fusion Tails for the Recovery and Purification of Recombinant Proteins," Protein Expr. Purif. 2:95-107 (1991).

X. Related Proceedings Appendix

No decisions have been rendered by a court or the Board in any proceedings identified in the *Related Appeals and Interferences* section (Section II).

Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 3372–3376, May 1988
Biochemistry

The α and β chains of avian retrovirus reverse transcriptase independently expressed in *Escherichia coli*: Characterization of enzymatic activities

(reverse transcriptase/bacterial expression)

DANIEL A. SOLTIS* AND ANNA MARIE SKALKA*†

Department of Molecular Oncology, Roche Institute of Molecular Biology, Roche Research Center, Nutley, New Jersey 07110

Communicated by John J. Burns, January 19, 1988

ABSTRACT Reverse transcriptase of the avian sarcoma and leukosis retroviruses is a heterodimer composed of a 63-kDa α and a 95-kDa β polypeptide chain, both of which are encoded in the *pol* gene and are produced by proteolytic processing of a larger precursor. We previously constructed a bacterial expression clone of the entire *pol* coding region that produces a protein 4 kDa larger than the mature viral β subunit. By use of this clone and synthetic oligonucleotides to introduce stop codons, two derivatives have been constructed: one that directs synthesis of a protein equivalent to the mature β subunit and the other that directs synthesis of a protein equivalent to α subunit. Predicted amino acid sequences of these proteins differ from their viral counterparts only by an initiator methionine that was added to the N termini for expression in *Escherichia coli*. Both bacterially expressed proteins exhibit reverse transcriptase activity and appear to function as homodimers. The properties of these proteins resemble those of the viral reverse transcriptase heterodimer; however, the bacterially produced α dimer protein could be distinguished from the other proteins by its increased sensitivity to heat inactivation, which also has been reported for the corresponding viral product. These results show that correct folding and expression of enzymatic function does not require formation of a precursor. The α and β clones provide a convenient source of individual *pol* gene products for further evaluation of their roles in the synthesis and integration of retroviral DNA.

The products of the *pol* gene of RNA tumor viruses are required for synthesis of viral DNA and its integration into the host genome (1, 2). Initially, *pol* is expressed together with *gag* as a fusion protein. This precursor is then processed by the viral *gag*- or *pol*-encoded protease to produce mature *gag* and *pol* proteins (3–5). In avian sarcoma and leukosis viruses (ASLV), the products of *pol* include three polypeptides that are formed by differential processing: a 95-kDa β chain, a 63-kDa α chain, and a phosphorylated 32-kDa chain designated pp32. The α sequence is identical to the N-terminal two-thirds of β, and the pp32 sequence is identical to the remaining C-terminal one-third of β (6, 7). The major form of the ASLV reverse transcriptase (RT) in virions is a heterodimer consisting of an α and β chain. The heterodimer possesses DNA polymerase activity that can copy RNA or DNA, RNase H activities, DNA-RNA unwinding activity, and a sequence-specific Mn^{2+}-dependent DNA endonuclease (1, 8). The pp32 chain has Mn^{2+}- and Mg^{2+}-dependent DNA endonuclease activity (9).

Efforts to evaluate the role of these various activities in retroviral DNA synthesis and integration are hampered somewhat by the fact that the *pol*-derived proteins do not normally exist in a soluble, active form in infected cells. The *gag–pol* fusion precursor appears to lack RT activity (10, 11), and the processed proteins, produced during or shortly after budding, are sequestered within the virion core. To circumvent this problem, we have prepared clones that express the mature proteins directly. In this study, we describe the construction of separate expression clones that encode the α or β subunits of ASLV *pol*, the partial purification of these proteins from bacteria, and a characterization of their enzymatic properties.

MATERIALS AND METHODS

Enzymes and Biochemicals. Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs. Calf intestine alkaline phosphatase and *Escherichia coli* DNA polymerase I Klenow fragment were obtained from Boehringer Mannheim. Deoxynucleoside [α-^{32}P]triphosphates and ^{125}I-labeled protein G were purchased from Amersham. Affinity-purified goat anti-rabbit IgG and rabbit anti-goat IgG were obtained from Cooper Biomedical. A poly(A)-selected mRNA preparation (mostly 1.2–6.0 kilobases in length) from murine lymphoma cell line 49.1, was the gift of U. Gubler of the Department of Molecular Genetics, Hoffmann–La Roche, and purified avian myeloblastosis virus reverse transcriptase (AMV-RT) was the gift of J. Leis of the Department of Biochemistry, Case Western Reserve University.

Construction of pRC23-p95. An expression clone, pRC23-p95, which encodes protein $p95^{pol\beta}$ (β subunit of RT) was constructed from pFA1-RT99 (12) by introduction of a synthetic double-stranded DNA fragment that contained a sequence coding for the C terminus of the β subunit (13) followed by a termination codon. The fragment was designed to have unpaired termini complementary to a *Ban* II restriction site for insertion in the appropriate location near the end of the *pol* gene (Fig. 1). All cloning steps including restriction digests, ligations, and transformations were by standard procedures (14, 15). The new construct, called pRC23-p95, was then introduced into MC1061 and RR1 strains of *E. coli* that contained the plasmid pRK248cIts, which expresses a temperature-sensitive λcI repressor.

Construction of pRC23-p63. A plasmid, pRC23-p63, which expresses $p63^{pol\alpha}$ (equivalent to the α subunit of ASLV-RT) from the λP_L promoter was constructed by oligonucleotide-directed site-specific mutagenesis (16) of pFA1-RT99. The oligonucleotide was homologous to coding sequences of the C terminus of α subunit and contained a termination codon plus additional sequences from another reading frame down-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ASLV, avian sarcoma and leukosis viruses; RT, reverse transcriptase; AMV-RT, avian myeloblastosis virus RT; ASLV-RT, ASLV RT.

*Present address: Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111.
†To whom reprint requests should be addressed.

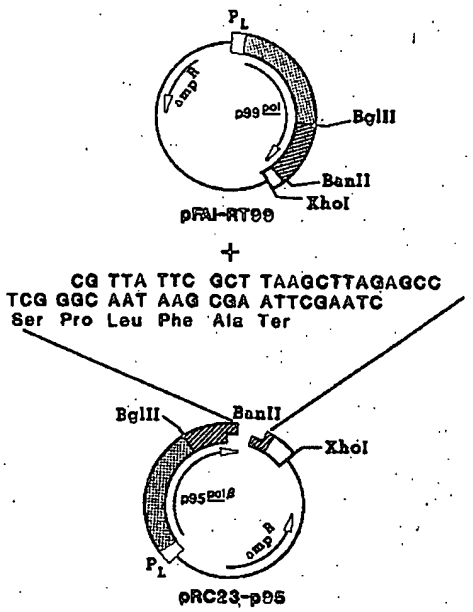


FIG. 1. Construction of $p95^{pol\beta}$ expression vector. A termination codon was inserted into pFA1-RT99 using synthetic oligonucleotides as described. The resulting plasmid, pRC23-p95, directs the expression of $p95^{pol\beta}$, a protein equivalent to the β subunit of ASLV RT. Sequences of the two synthetic oligodeoxynucleotides and the predicted amino acids they encode are shown. Ter, termination codon for $p95^{pol\beta}$; □, P_L promoter; ▨, α coding region of *pol* gene; ▩, endonuclease region of *pol*; and □, viral envelope sequences.

stream in *pol* (Fig. 2). By hybridizing the oligonucleotide to pFA1-RT99 DNA, which was manipulated so that it contained a single-stranded gap in the *pol* coding sequences, the termination codon was brought into frame at the end of the α subunit coding sequences (13, 17), and the intervening 68 bases were looped-out. After repairing the single-stranded gap with the Klenow fragment of *E. coli* DNA polymerase I, the plasmid containing a region of heteroduplex was introduced into *E. coli*, and a clone was obtained from which only $p63^{pol\alpha}$ was expressed (Fig. 2).

Enzyme Preparation. The *E. coli* strain MC1061 (18) containing the expression vector for $p95^{pol\beta}$ (pRC23-p95) or $p63^{pol\alpha}$ (pRC23-p63) and the plasmid encoding the temperature-sensitive λcI repressor (pRK248cIts) were grown at 30°C in M9 medium supplemented with 0.2% glucose, 0.2% Casamino acids, and ampicillin at 50 μg/ml, to an A_{600} of 1.0.

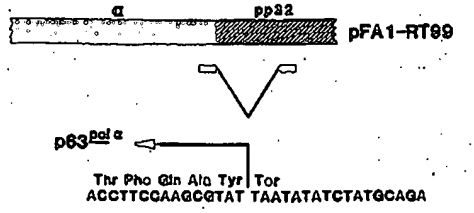


FIG. 2. Construction of $p63^{pol\alpha}$ expression vector. Using oligonucleotide-directed site-specific mutagenesis of pFA1-RT99, a termination codon was brought into frame at the end of the α coding sequences of the *pol* gene, as described. Plasmid pRC23-p63 directs the expression of $p63^{pol\alpha}$, which is equivalent to the α subunit of ASLV-RT. Sequence of the synthetic oligodeoxynucleotide used and the predicted amino acid sequence at the C terminus of $p63^{pol\alpha}$ are shown. Ter, termination codon for $p63^{pol\alpha}$; ▨, α coding region of *pol* gene; and ▩, endonuclease region of *pol*.

Protein expression was induced by shifting the cultures to 42°C for 3 hr. Bacteria were harvested by centrifugation at 6000 × *g* for 10 min at 4°C, and the pellets were stored at −20°C.

The bacterial pellets (2–3 g) were resuspended in 50 mM Tris·HCl, pH 7.5/5 mM dithiothreitol/10% (wt/vol) sucrose, containing lysozyme (2 mg/ml) at a concentration of 0.4 g of cells per ml of solution. The suspension was incubated for 30 min at 4°C. EDTA, pH 8.0, was added to a final concentration of 5 mM, and the mixture was incubated for 10 min at 4°C. An equal volume of 50 mM Tris·HCl, pH 7.5/5 mM dithiothreitol/1% Brij-58 was added, and the suspension was incubated for 30 min at 4°C. This was then frozen in a dry ice/ethanol bath and thawed at 37°C to increase the bacterial lysis. The bacterial debris was removed by centrifugation for 50 min at 80,000 × *g* and 4°C, the supernatant was removed, and a sample was taken for assay of RT activity. Streptomycin sulfate was then added slowly to the supernatant, with stirring, to a final concentration of 0.83%, and the resulting suspension was stirred at 4°C for 30 min. The precipitate was collected by centrifugation at 30,000 × *g* and 4°C for 30 min, and the pellet was resuspended in 10 ml of 20 mM Tris·HCl, pH 7.5/1 mM dithiothreitol/0.1 mM EDTA/10% glycerol. Ammonium sulfate was then added, with stirring, to 70% saturation, and the mixture was kept on ice for 3–5 hr. The precipitate was collected by centrifugation for 40 min at 45,000 × *g* and 2°C. The pellet was resuspended in 3 ml of 20 mM potassium phosphate, pH 7.5/2 mM dithiothreitol/20 mM KCl/10% glycerol (buffer A) and dialyzed at 4°C twice in 1 liter each time of buffer A. The dialyzed samples were then loaded onto 1-ml hydroxylapatite (Bio-Gel HPT; Bio-Rad) columns that had been equilibrated in 20 mM potassium phosphate, pH 7.5/2 mM dithiothreitol/10% glycerol (buffer B). Columns were washed in 2 ml of buffer B, and the enzymatic activity was eluted using a step gradient in buffer B in which the KCl concentration was increased from 0.1 M to 0.3 M and finally to 1.0 M. Fractions were then assayed for RT activity. For $p95^{pol\beta}$ the activity eluted in the 1 M KCl fractions, and for $p63^{pol\alpha}$, the activity eluted in the 0.3 M KCl fractions. Fractions containing like enzymatic activity were pooled and dialyzed in buffer A. The dialyzed hydroxylapatite-column pools were then assayed for nuclease and background polymerase activities that would interfere with RT assays by doing mixing experiments using AMV-RT as described in *Results*. The $p63^{pol\alpha}$ pool was essentially free of interfering activities and was stored at 4°C; however, the $p95^{pol\beta}$ pool still contained substantial amounts of these activities and required further purification.

The dialyzed $p95^{pol\beta}$ hydroxylapatite pool was loaded onto a 1-ml poly(uridylic acid)-agarose (Pharmacia) column equilibrated in 20 mM potassium phosphate, pH 7.5/2 mM dithiothreitol/20 mM KCl/20% (vol/vol) glycerol (buffer C). The column was then washed with 2 ml of buffer C, and the enzymatic activity was eluted using a step gradient in buffer C in which the KCl concentration was increased from 0.1 M to 0.3 M, to 1.0 M, and finally to 2.0 M. Fractions were then assayed for RT activity and the $p95^{pol\beta}$ was eluted in the 0.3 M and 1.0 M KCl-containing fractions. These fractions were pooled and diluted in buffer C before concentrating by reloading onto a 1-ml hydroxylapatite column as previously described. The enzymatic activity was eluted in a small volume of buffer B containing 1 M KCl. Fractions containing RT activity were pooled and then dialyzed in buffer A at 4°C. This preparation was found to be essentially free of activities that interfered with the RT assay and was stored at 4°C.

RT Assays. RT activity was determined by measuring the addition of [α-^{32}P]dGTP to an oligo(deoxyguanidylic acid)[oligo(dG)] primer hybridized to poly(ribocytidylic acid-)[poly(rC)] with a modification of a published procedure (19). Concentrations of the substrates used in the assays were 260 μM poly(rC), 98 μM oligo(dG) [12–18 nucleotides (nt) in

length] and 10 μM [α-^{32}P]dGTP (ca. 20 Ci/mmol; 1 Ci = 37 GBq); all concentrations are in nucleotides. The enzyme fraction was added to the reaction mixture, which was then incubated for 30 min at 37°C. Samples were spotted on DE-81 paper; the paper was washed in 0.3 M sodium chloride/30 mM sodium citrate and dried, and the radioactivity was measured by Cherenkov counting in a liquid scintillation counter.

Product Analysis. The reactions were monitored as described above. They contained 370 μM poly(A)-selected mRNA from a murine lymphoma cell line, 12.5 μM oligo(deoxythymidylic acid) (12–18 nt), 10 μM dATP, 10 μM dTTP, 1 μM [α-^{32}P]dCTP (320 Ci/mmol), 1 μM [α-^{32}P]dGTP (320 Ci/mmol), and actinomycin D at 50 μg/ml. These reactions were incubated at 37°C for 30 min. Samples from reaction mixtures were dried by lyophilization, and the residues were resuspended in 7 M urea/89 mM Tris borate, pH 8.3/1 mM EDTA, pH 8.0/0.01% (wt/vol) bromphenol blue/0.01% (wt/vol) xylene cylanol. The samples were then boiled for 5 min and loaded onto a 26 × 14 cm sequencing gel containing 7.5% acrylamide, 0.2% bisacrylamide, 7 M urea, 89 mM Tris borate (pH 8.3), and 1 mM EDTA. Electrophoresis was done for 3 1/2 hr at 600 V, and an autoradiograph of the gel was then prepared using Kodak XAR-5 film.

NaDodSO$_4$/PAGE and Immunoblotting. Protein samples were boiled for 5–10 min and then fractionated by electrophoresis in a NaDodSO$_4$/9% polyacrylamide gel as described by Laemmli (20). The proteins were then stained with Coomassie brilliant blue or transferred from the gel to a 0.2-μm nitrocellulose filter (21). The filter was blocked and treated as previously described using as the primary antibody an affinity-purified rabbit polyclonal antipeptide antibody directed against a sequence in the α portion of the *pol* gene (12). An autoradiograph of the filter was prepared by exposing Kodak XAR-5 film to the filter at −70°C.

Glycerol Density Gradient Sedimentation. Samples (150 μl) that included bovine serum albumin at 267 μg/ml (fraction V), which served as an internal marker, were layered on 10–30% linear glycerol gradients containing 50 mM potassium phosphate (pH 7.5), 2 mM dithiothreitol, 0.1% Nonidet P-40, and 300 mM KCl. After centrifugation for 14 hr at 200,000 × *g* and 5°C in a Beckman SW55 rotor, the gradients were fractionated by collecting drops from a 21-gauge needle pushed through the bottom of the centrifuge tube. Fractions were assayed for RT activity and/or analyzed by NaDodSO$_4$/PAGE.

RESULTS

Cloning and Expression of ASLV *pol*-Related Gene Products in *E. coli*. With the plasmid pFA1-RT99 (12), which expresses the entire *pol* gene coding region, two additional bacterial expression clones were constructed to obtain proteins with the sequence of the mature α and β subunits of ASLV-RT. Approximately 90% of the p63$^{pol\alpha}$ and p95$^{pol\beta}$ proteins produced in *E. coli* bacteria was present in insoluble inclusion bodies, as had been seen previously with the p99^{pol} protein (12). Attempts to solubilize this material in chaotropic agents such as urea and guanidinium hydrochloride or high levels of various detergents were successful; however, this solubilized protein lacked RT activity. Therefore, for further analyses, we focused on the enzymatically active soluble fraction in the lysates. Incorporation of dGTP with this fraction depended on the presence of poly(rC), as was the case with AMV-RT. No incorporation was detected in lysates of bacteria that contained pRC23 (22) with no insert or that expressed another viral protein (data not shown).

Partial Purification of the *pol*-Derived Proteins from *E. coli*. The *pol*-derived proteins in the soluble fraction of the lysate were purified further to remove other enzymatic activities, including nucleases and a poly(rC)-independent DNA polymerase, which interfered with the RT assay.

Table 1. Partial purification of p95$^{pol\beta}$ and p63$^{pol\alpha}$ from *E. coli*

	RT unit/mg of protein	
	p95$^{pol\beta}$	p63$^{pol\alpha}$
Lysate	6.0×10^{-3}	5.9×10^{-4}
Resuspended $(NH_4)_2SO_4$ pellet	5.8×10^{-2}	7.6×10^{-4}
Hydroxylapatite column pool	8.2×10^{-1}	2.1×10^{-1}
Poly(U)-agarose column pool	2.7	

RT assays of fractions from the partial purification of p95$^{pol\beta}$ and p63$^{pol\alpha}$ were done as described. A unit of RT activity is defined as the amount of enzyme that incorporates 1 nmol of dNTP in 30 min at 37°C.

Table 1 summarizes the results from partial purification of p95$^{pol\beta}$ and p63$^{pol\alpha}$. Both the poly(U)-agarose fraction of p95$^{pol\beta}$ and the hydroxylapatite-column fraction of p63$^{pol\alpha}$ were relatively free of interfering activities, as shown by the lack of RT inhibition in a mixing experiment that contained purified AMV-RT and a sample of the partially purified proteins (data not shown). These preparations were also free of poly(rC)-independent DNA polymerase activity. Specific activities of the p95$^{pol\beta}$ and p63$^{pol\alpha}$ proteins increased ≈1000-fold during purification. However, due to interfering activities present in early fractions that had made accurate quantitation of RT activity impossible, the increases seen in specific activity probably do not correspond to a similar increase in purity of the p95$^{pol\beta}$ and p63$^{pol\alpha}$ proteins. Final specific activities are still lower by factors of 10,000 to 100,000 than that of purified AMV-RT, and protein analyses indicated that both p63$^{pol\alpha}$ and p95$^{pol\beta}$ were minor constituents in these preparations. By use of additional column chromatography steps, small amounts of a more highly purified fraction of the bacterially produced protein (≈20% pure) were obtained. Specific activity of this preparation was within a factor of ten that of viral RT; however, insufficient protein was obtained for further characterization.

Fig. 3 shows the results of immunoblot analysis of the proteins in samples from various steps noted in Table 1. In all samples, the bacterially produced p95$^{pol\beta}$ protein comigrated with the β subunit of AMV-RT, and the p63$^{pol\alpha}$ protein comigrated with the α subunit, indicating that the termination codons were inserted in the proper positions in these clones and that they functioned correctly. At least eight prominent smaller *pol*-related proteins in the crude extracts were removed during purification. These proteins, which ranged in size from ≈73 to 35 kDa, may be immunologically related *E. coli* bands, products of proteolytic cleavage in *E. coli*, or the result of premature termination or incorrect initiation during synthesis. Both p95$^{pol\beta}$ and p63$^{pol\alpha}$ proteins appeared stable during purification, and final preparations contain only one major *pol*-related protein.

RT Assays and Product Characterization. Properties of the partially purified p95$^{pol\beta}$ and p63$^{pol\alpha}$ proteins were compared with those of AMV-RT, and the results are summarized in Table 2. Without template, poly(rC), no incorporation of

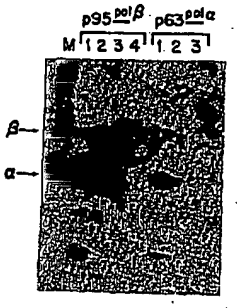


FIG. 3. Partial purification of p95$^{pol\beta}$ and p63$^{pol\alpha}$. An immunoblot of samples from the purification using α-specific antiserum as described. Lane M, AMV-RT; p95$^{pol\beta}$ lanes: 1, lysate soluble fraction (10 μg of protein); 2, resuspended $(NH_4)_2SO_4$ pellet (5.2 μg); 3, hydroxylapatite column pool (90 μg); 4, poly(U)-agarose column pool (19 μg); p63$^{pol\alpha}$ lanes: 1, lysate soluble fraction (4.4 μg); 2, resuspended $(NH_4)_2SO_4$ pellet (6.6 μg); 3, hydroxylapatite column pool (95 μg).

Table 2. Comparison of properties of $p95^{pol\beta}$, $p63^{pol\alpha}$, and AMV-RT

	RT activity, %		
	AMV-RT	$p95^{pol\beta}$	$p63^{pol\alpha}$
Control	100	100	100
Omit poly(rC)	<0.01	<0.02	<0.01
10 mM *N*-ethylmaleimide (10 min at 0°C)	<0.5	<0.5	<0.5
Heated (15 min at 44°C)	84	95	10
Goat polyclonal antiserum (60 min at 15°C)	<0.1	<0.1	<0.1
Normal goat serum (60 min at 15°C)	93	89	74

RT assays were done as described after treating the proteins as indicated. *N*-ethylmaleimide was incubated with samples of the proteins at pH 8.0 for 10 min on ice. The reactions were stopped by the addition of an equal volume of RT assay buffer containing 50 mM dithiothreitol. Thermal stability of the proteins was analyzed by incubation at 44°C for 15 min in RT assay buffer containing 260 μM poly (rC) and 98 μM oligo(dG), but containing no [α-^{32}P]dGTP. Reactions were stopped by adding an equal volume of the same buffer to the reaction and placing the mixture on ice. RT assays were done on these samples by adding 1/2 volume of the buffer described above containing 30 μM [α-^{32}P]dGTP (*ca.* 20 Ci/mmol) to make a complete reaction mixture. Incubations were done, and samples were treated as described. The 100% values for control samples correspond to the incorporation of from 0.1 to 5 pmol of deoxynucleoside triphosphate by the enzyme indicated after incubation under the appropriate control conditions.

[α-^{32}P]dGTP was detected with any enzyme preparations. All three RT activities could be inactivated by incubation with a goat polyclonal antiserum prepared against purified AMV-RT, but these activities were essentially unaffected by normal goat serum. The RT activity of AMV-RT was completely inactivated by incubation with 10 mM *N*-ethylmaleimide for 10 min on ice, and the activities of $p95^{pol\beta}$ and $p63^{pol\alpha}$ were similarly sensitive to this compound. On the other hand, AMV-RT is reported to be relatively resistant to heat inactivation (23). We found that in the presence of template and primer [oligo(dG)-primed poly(rC)], but with no dGTP substrate, the RT activities of AMV-RT and $p95^{pol\beta}$ were essentially unaffected by incubation at 44°C for 15 min. The activity of $p63^{pol\alpha}$, however, was reduced by a factor of 10 under the same conditions.

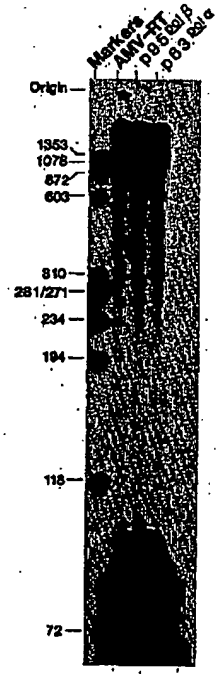


FIG. 4. Product analysis. RT assays were done using a poly(A)-selected mRNA primed with oligo(dT). Reaction products were analyzed as described. Marker lane, ^{32}P-labeled *Hae* III digest of bacteriophage φX174 DNA (numbers indicate length of fragments); lane AMV-RT, products synthesized by 1.4×10^{-3} units of AMV-RT; lane $p95^{pol\beta}$, products synthesized by 0.6×10^{-3} unit of $p95^{pol\beta}$; lane $p63^{pol\alpha}$, products synthesized by 0.2×10^{-3} unit of $p63^{pol\alpha}$.

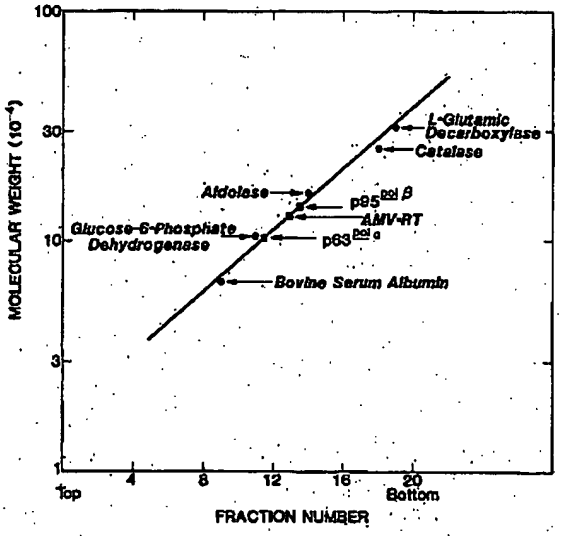


FIG. 5. Glycerol-gradient sedimentation of $p63^{pol\alpha}$, $p95^{pol\beta}$, and AMV-RT. Sedimentation analysis was done as described; 8.5 units of AMV-RT, 0.10 units of $p95^{pol\beta}$, and 0.02 units of $p63^{pol\alpha}$ were applied separately to each of three gradients. Positions of the M_r markers and of $p63^{pol\alpha}$, $p95^{pol\beta}$, and AMV-RT sedimented in parallel gradients are indicated. L-glutamic decarboxylase, M_r 310,000; catalase, M_r 250,000; aldolase, M_r 160,000; glucose-6-phosphate dehydrogenase, M_r 104,000; and bovine serum albumin, M_r 66,000.

The products of reverse transcription by AMV-RT, $p95^{pol\beta}$, and $p63^{pol\alpha}$ on a poly(A)-selected mRNA template primed with oligo(dT) are shown in Fig. 4. The reactions were run in the presence of actinomycin D (50 μg/ml) to prevent second-strand synthesis, and the products were analyzed by PAGE. With all three proteins, two populations of products were seen. One consisted of very long molecules, most of which exceeded 1.3 kilobases. The second, representing the majority of incorporated nucleotides, were <20 nt. When actinomycin D was omitted, the products synthesized by each of the three proteins were again similar. However, in this case, most products were >100 nt and only a small percentage were <30 nt (data not shown).

Glycerol Gradient Sedimentation of Recombinant Proteins. Samples of partially purified $p95^{pol\beta}$ and $p63^{pol\alpha}$ and purified AMV-RT were sedimented in glycerol gradients under nondenaturing conditions to determine their quaternary (multimeric) structures. All three proteins sedimented in single peaks as determined by RT assays of the gradient fractions. Fig. 5 shows that the AMV-RT sedimented as a M_r of ≈125,000 with respect to protein standards. This M_r is somewhat lower than expected from the calculated M_r of the AMV-RT αβ heterodimer (24, 25). The $p95^{pol\beta}$ protein sedimented as 140,000, and the $p63^{pol\alpha}$ protein sedimented as 105,000. Thus, both sediment as homodimers of the subunits that compose the AMV-RT heterodimer.

DISCUSSION

We expressed three different products of the ASLV *pol* gene in *E. coli* and compared some of their properties after partial purification. A full-length product, p99^{pol}, that includes 4 kDa of C-terminal sequences not present in the mature ASLV β

subunit, had been described (12). Our constructs express a 95-kDa protein the amino acid sequence of which is equivalent to the β subunit of ASLV-RT, and a 63-kDa protein equivalent to the α subunit. We have not determined if the N-terminal methionine residue, required for expression, was removed from these proteins as occurred with the *pol*-derived $p36^{pol}$ protein expressed in *E. coli* (12). However, unlike previous RT constructions that utilized *pol* sequences from murine leukemia virus (MLV) (26–29), these proteins were not expressed as fusions with other nonrelated sequences. Thus, we expect their properties to be similar, if not identical, to those of the corresponding viral protein subunits.

The level of expression of $p95^{pol\beta}$ and $p63^{pol\alpha}$ in *E. coli* is about 1% of total cellular protein. However, in both cases, ≈90% of the protein was sequestered in insoluble inclusion bodies, and attempts to recover enzymatically active preparations from this material were unsuccessful. Enzymatically active protein was detected in the soluble fraction of the bacterial lysates, but it was difficult to quantitatively determine the amount of active protein present due to contaminating activities that interfered with the RT assay. Other *E. coli* host/vector combinations or other microbial expression systems might provide a higher yield of soluble active protein. To verify that our constructs produced proteins with the expected activities, we partially purified the enzymatically active soluble portions until free of components that could interfere with the RT assay. We also attempted to remove most, if not all, of the smaller *pol*-related proteins from preparations, so that the properties measured would reflect those of the 95-kDa (β) and 63-kDa (α) products. The chromatographic procedures that proved successful are, in fact, similar to those that have been used to purify the analogous viral proteins (30).

Glycerol-gradient sedimentation analyses of the partially purified $p95^{pol\beta}$ and $p63^{pol\alpha}$ proteins suggest that the enzymatically active forms of both proteins are homodimers. Enzymatically active β dimers have been purified from ASLV particles, in which they constitute a variable percentage of the total RT activity (30). In virions the α form may exist as either a monomer (30) or a homodimer (25), and what factors effect the subunit structure are unclear.

The properties of the bacterially produced proteins that were soluble and active were very similar to those of the viral RT. All three were completely inactivated after incubating with *N*-ethylmaleimide or with a goat polyclonal antiserum directed against purified AMV-RT, and all three produced similar cDNAs from an mRNA template. The $p63^{pol\alpha}$ protein could, however, be distinguished from AMV-RT and $p95^{pol\beta}$ by its increased sensitivity to heat inactivation. A similar observation has been reported for the α subunit purified from ASLV virions (30).

A major purpose of this study was to determine whether the α ($p63^{pol\alpha}$) and β ($p95^{pol\beta}$) constructs could produce active proteins. Because these are large eukaryotic proteins, it was not known whether prokaryotic cells would provide a suitable test system. During viral infection, the ASLV *pol*-related proteins are produced as polyprotein precursors that are processed, presumably after aggregation, and then further modified by phosphorylation. It was not known whether either one of these processes would effect the final conformation or activity of the mature proteins. Our results show that the α and β chains of ASLV-RT can be expressed directly in *E. coli* and that the properties of the soluble, enzymatically active proteins resemble those of the viral proteins. Although technical difficulties now limit the system as a source of large amounts of protein, sufficient enzymatic activity exists for these clones to be useful in mutagenesis studies. The coding regions in these clones can also be transferred to a variety of eukaryotic expression vectors for use in genetic complementation studies and other tests of their function.

We thank M. Poonian of the Department of Molecular Genetics, Hoffmann–La Roche, for synthesis of the oligonucleotides used in this study. This work was supported by National Institutes of Health Grants CA-06927, RR-05539, a grant from the Pew Charitable Trust, and also by an appropriation from the Commonwealth of Pennsylvania.

1. Varmus, H. E. & Swanstrom, R. (1982) in *RNA Tumor Viruses*, eds. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 369–512.
2. Panganiban, A. T. & Temin, H. M. (1984) *Cell* 36, 673–679.
3. Eisenman, R. N., Mason, W. S. & Linial, M. (1980) *J. Virol.* 36, 62–78.
4. Moelling, K., Scott, A., Dittmar, K. E. J. & Owada, M. (1980) *J. Virol.* 33, 680–688.
5. Oppermann, H., Bishop, J. M., Varmus, H. L. & Levintow, R. (1977) *Cell* 12, 993–1005.
6. Rho, H. M., Grandgenett, D. P. & Green, M. (1975) *J. Biol. Chem.* 250, 5278–5280.
7. Schiff, R. D. & Grandgenett, D. P. (1978) *J. Virol.* 28, 279–291.
8. Golomb, M. & Grandgenett, D. P. (1979) *J. Biol. Chem.* 254, 1606–1613.
9. Grandgenett, D. P., Vora, A. C. & Schiff, R. D. (1978) *Virology* 89, 119–132.
10. Panet, A. & Baltimore, D. (1987) *J. Virol.* 61, 1756–1760.
11. Witte, O. N. & Baltimore, D. (1978) *J. Virol.* 26, 750–761.
12. Alexander, F., Leis, J., Soltis, D. A., Crowl, R. M., Danho, W., Poonian, M. S., Pan, Y. E. & Skalka, A. M. (1987) *J. Virol.* 61, 534–542.
13. Grandgenett, D., Quinn, T., Hippenmeyer, P. J. & Oroszlan, S. (1985) *J. Biol. Chem.* 260, 8243–8249.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
15. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *Advanced Bacterial Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Morinaga, Y., Franceschini, T., Inouye, S. & Inouye, M. (1984) *Biotechnology* 2, 636–639.
17. Copeland, T. D., Grandgenett, D. P. & Oroszlan, S. (1980) *J. Virol.* 36, 115–119.
18. Casadaban, M. J. & Cohen, S. W. (1980) *J. Mol. Biol.* 138, 179–207.
19. Goff, S., Traktman, P. & Baltimore, D. (1981) *J. Virol.* 38, 239–248.
20. Laemmli, U. K. (1970) *Nature (London)* 277, 680–685.
21. Khyse-Andersen, J. (1984) *J. Biochem. Biophys. Methods* 10, 203–209.
22. Crowl, R., Seamans, C., Lomedico, P. & McAndrew, S. (1985) *Gene* 38, 31–38.
23. Verma, I. M., Mason, E. S., Drost, S. D. & Baltimore, D. (1974) *Nature (London)* 251, 27–31.
24. Hurwitz, J. & Leis, J. P. (1972) *J. Virol.* 9, 116–129.
25. Grandgenett, D. R., Gerard, G. F. & Green, M. (1973) *Proc. Natl. Acad. Sci. USA* 70, 230–234.
26. Tanese, N., Roth, M. & Goff, S. P. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4944–4948.
27. Kotewicz, M. L., D'Allessio, J. M., Driftmier, K. M., Blodgett, K. P. & Gerard, G. F. (1985) *Gene* 35, 249–258.
28. Roth, M. J., Tanese, N. & Goff, S. P. (1985) *J. Biol. Chem.* 260, 9326–9335.
29. Gerard, G. F., D'Alessio, J. M., Kotewicz, M. L. & Noon, M. C. (1986) *DNA* 5, 271–279.
30. Hizi, A. & Joklik, W. K. (1977) *J. Biol. Chem.* 252, 2281–2289.

Proceedings
OF THE
National Academy of Sciences
OF THE UNITED STATES OF AMERICA

May 1988 | Volume 85, Number 10 | pp. 3271–3660

Table of Contents

Contents

PROTEIN EXPRESSION AND PURIFICATION 4, 298–303 (1993)

Expression and Purification of the HIV-1 Reverse Transcriptase Using the Baculovirus Expression Vector System[1]

Simon Kawa,* Amalandra Kumar,† John S. Smith,* S. Patricia Becerra,† William A. Beard,† Samuel H. Wilson,† and E. Brad Thompson*,[2]

*Department of Human Biological Chemistry and Genetics, BSB. Room 603, The University of Texas Medical Branch, Galveston, Texas 77555-0645; †Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892; and Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, Texas 77555-0851

Received October 28, 1992, and in revised form February 26, 1993

We have successfully expressed and purified the human immunodeficiency virus type-1 reverse transcriptase (RT) using the baculovirus expression vector system. This expression system provides a eukaryotic environment in which post-translational modifications of foreign gene products can occur. After infection with recombinant virus, Western blot analysis confirmed the presence of an immunoreactive polypeptide of approximately 66 kDa from insect Sf9 cell lysates. RT was then purified from crude extracts of baculovirus-infected Sf9 cells; SDS–PAGE analysis of fractions obtain from partial purification showed that in contrast to the *Escherichia coli*-expressed RT, the baculovirus-expressed RT corresponded to a doublet of peptides at approximately 66 kDa. Further purification of the protein resulted in a p66 protein, judged to be more than 90% pure by SDS–PAGE and Coomassie blue stain. Following purification, the baculovirus derived RT had specific activity for DNA polymerase similar to that of the *E. coli*-derived RT. Therefore, RT purified from Sf9 cells appears to be suitable for structure–function studies of this enzyme.

Infection of cells by the human immunodeficiency virus type-1 (HIV-1) results in the establishment of a DNA copy of the viral RNA genome. This RNA-directed DNA synthesis is catalyzed by the enzyme reverse transcriptase (RT), encoded by the 5′ end of the *pol* gene (1). RT activity in virions is associated with the presence of two polypeptides, p66 (M_r ~66,000) and p51 (M_r ~51,000), which share a common N-terminus; p51 is derived from p66 by a proteolytic cleavage (2,3). The region of the *pol* gene encoding RT has been cloned, and expression of p66 has been achieved in bacteria by several laboratories (4–11). This expression leads to isolation of smaller amounts of a 66-kDa homodimer and larger amounts of a p66/p51 heterodimer form of the enzyme including the carboxyl-terminally truncated 51-kDa peptide. The latter arises via proteolysis of the 66-kDa protein which leads to a small increase in RT activity.

Despite these advances, many aspects of the HIV-1 RT structure and function remain unknown. This has necessitated the search of expression systems to study the properties of the enzyme and the proteolysis that leads to the 51-kDa form. We report that we have successfully expressed and purified RT derived from HXB2 proviral DNA (12) in Sf9 cells using the baculovirus expression vector system (BEVS).

The BEVS is based on the life cycle of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) in Sf9 cells. Briefly, late in the viral life cycle, a polyhedrin protein accumulates to high levels and accounts for 50–70% of the total cell protein. The polyhedrin is nonessential for viral infection or replication. Foreign genes can be inserted into AcMNPV and expressed to high levels under control of the polyhedrin promoter. This expression system provides a eukaryotic environment in which post-translational modifications required for the biological activity of some eukaryotic proteins occur. Infection of Sf9 cells with the recombinant virus results in the expression of RT approximately 48 h post-infection. Western blot analysis using a monoclonal an-

[1] This work was supported by NIH Grant 5R01 DK41058 to E. Brad Thompson and by the National Institutes of Health Intramural AIDS Targeted Antiviral Program to Samuel H. Wilson.

[2] To whom correspondence should be addressed.

1046-5928/93 $5.00

tibody to RT confirmed the presence of the RT polypeptide from infected cell lysates, which exhibited electrophoretic mobility similar to that of the *Escherichia coli*-expressed RT. Following purification the baculovirus-derived enzyme had specific activity for DNA polymerase similar to that of the *E. coli*-derived RT.

MATERIALS AND METHODS

Buffers

All buffers used during protein purification contained the following protease inhibitors: 1 mM PMSF, 10 mM $Na_2S_2O_5$, and 1 μg/ml pepstatin A. The lysis buffer contained 25 mM Tris–HCl, pH 8.0, 200 mM NaCl, 20% glycerol, 1% Triton X-100, 2 mM DTT, and 1 mM EDTA. Buffer A contained 50 mM Tris–HCl, pH 7.0, 1 mM DTT, 1 mM EDTA, 0.1% NP-40, and 10% glycerol. TBS Buffer was 20 mM Tris–HCl, pH 7.5, and 150 mM NaCl and TBST was TBS containing 0.05% Tween 20.

Cell Cultures

Sf9 cells were cultured at 27°C in TNM-FH medium supplemented with 10% fetal bovine serum (FBS) as described (13). TNM-FH medium contained Grace's medium (GIBCO/BRL, Grand Island, NY), lactalbumin hydrolysate, and yeast extract. Sf9 cells are not anchorage dependent and can be transferred between monolayer and suspension without noticeable loss of viability or growth rate. Sf9 cells grown in monolayer cultures were infected with AcMNPV at a multiplicity of infection (moi) of 5 10. After 48 h postinfection, the medium containing the extracellular virus and the viral DNA was purified as described (13).

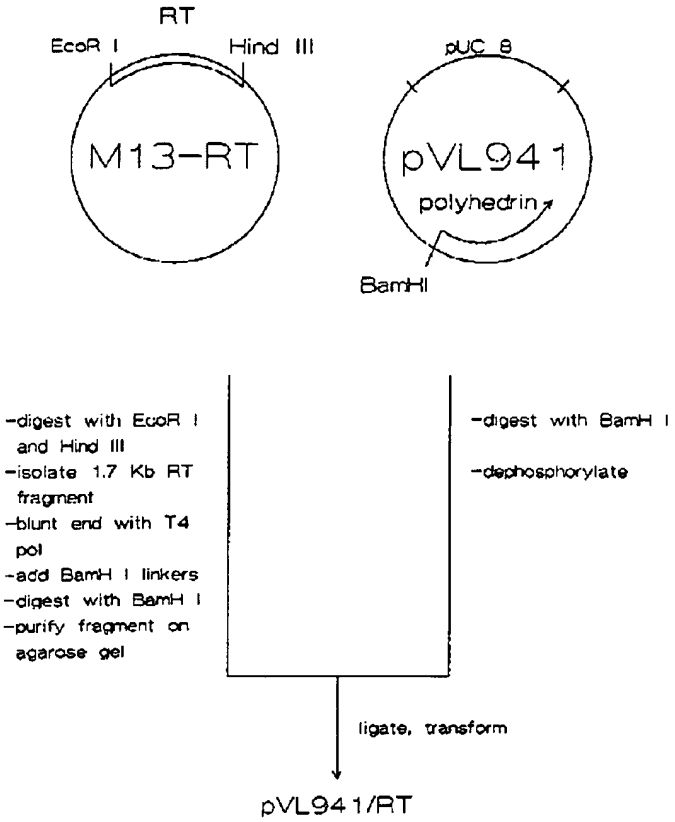


FIG. 1. Schematic diagram of the construction of the recombinant plasmid pVL941/RT. The RT cDNA fragment obtain by digesting M13-RT with *Eco*RI and *Hin*dIII was blunt-ended with T_4 polymerase and *Bam*HI linkers were added as described. The linked fragment was digested with *Bam*III to generate cohesive termini. pVL941 was also digested with *Bam*HI and dephosphorylated. pVL941 and RT cDNA fragments were ligated with T_4 ligase. The ligation mixture was used to transform competent DH5α cells.

Construction of PVL941/RT

To purify RT cDNA the plasmid M13-RT (11) derived from HXB2 proviral DNA (12) was digested completely with *Eco*RI and *Hin*dIII. The fragments were separated by electrophoresis on an 0.8% agarose gel. The 1.7-kb fragment containing RT cDNA was electroeluted and purified on a elutip column (Schleicher and Schuell, Inc., Keene, NH). The RT fragment was blunt-ended with T_4 polymerase. *Bam*HI linkers were added, and the linked fragment was digested with *Bam*HI to generate cohesive termini and inserted into the *Bam*HI site of PVL941 to generate the recombinant plasmid PVL941/RT (Fig. 1). The recombinant plasmid was transformed into competent DH5α cells, amplified, and purified as described (14).

Generation of Recombinant Virus

Cotransfection of PVL941/RT and AcMNPV into Sf9 cells was performed by calcium phosphate precipitation (13). Transfer of RT cDNA sequences from the recombinant plasmid to the AcMNPV genome occurs by homologous recombination. At 96 h postinfection, the medium containing a mixture of wild-type and recombinant viruses was collected. The recombinant virus was isolated by several rounds of plaque purification, due to the distinct plaque morphology of cells infected with the recombinant viruses (13).

Dot Blot Screening of RT Recombinant Plaques

Recombinant plaques were picked and resuspended in 100 μl TNM-FH + 10% FBS, then placed in a 96-well microtiter plate with 100 μl SF9 cells (1.5×10^5 cells/ml). After 4 days, supernatants were removed and transfered to a second 96-well plate. The Sf9 cells that were attached to the bottom of the wells were treated with 200 μl 0.5 N NaOH for 10 min and then with 20 μl 10 M ammonium acetate, mixed, and spotted on a Gene Screen Plus Nylon membrane (New England Nuclear,

Boston, MA) using a dot blot manifold apparatus (Bio-Rad Laboratories, Richmond, CA). Cell lysates were incubated on the membrane for 40 min at room temperature and then gentle suction was applied to the blotting apparatus. The membrane was removed and air dried. Prehybridization and hybridization were performed as described (14), using ^{32}P-labeled RT cDNA as a probe. The membrane was washed and exposed to a Kodak X Omat film for 15 h at −70°C.

Determination of RT Activity in Sf9 Cells

Sf9 cells in monolayer cultures were infected with the recombinant virus at a moi of 1–10. At 48 h postinfection cells were harvested by centrifugation at 1000*g* for 5 min and the cell pellets placed at −70°C. Frozen cells were thawed, resuspended in 500 μl TE, and centrifuged at 16,000*g* for 5 min. Thirty-five microliters of supernatant was assayed for incorporation of [^{3}H]thymidine 5'triphosphate into DNA fragments using a poly(rA)-oligo(dT)$_{12-18}$ template. The mixture was incubated for 1 h and the DNA-incorporated label was captured by adsorption on DE-81 filter paper (Whatman, Inc., Clifton, NJ).

Protein Purification

A total of 10 g of insect cell pellet was suspended in 100 ml lysis buffer. The cell suspension was probe-sonicated with four bursts, 30 s each, at 50 W with the tube immersed in an ice-water bath. The lysate was centrifuged at 5000*g* for 20 min to remove particulate material. The supernatant fraction was dialyzed against Buffer A containing 75 mM NaCl at 4°C for 3 h. The dialyzed extract was centrifuged at 10,000*g* for 15 min, and the supernatant was layered on a Q Sepharose fast flow column (30-ml bed volume, Pharmacia LKB Biotechnology, Inc., Piscataway, NJ) equilibrated with buffer A containing 75 mM NaCl. The flowthrough fraction was collected (Pool I).

Pool I was layered on a single-stranded DNA (ssDNA)–cellulose column (25-ml bed volume, Sigma Chemical Co., St. Louis, MO) equilibrated with Buffer A containing 75 mM NaCl. The column was washed with 1 column vol of Buffer A containing 75 mM NaCl, and p66 was eluted with a linear gradient of 75–600 mM NaCl in Buffer A (total volume 140 ml) at 0.5 ml/min. Fractions of 2 ml were collected. These fractions were analysed by SDS–PAGE and several fractions containing the peak of p66 were pooled (Pool II, 35 ml). Pool II was concentrated using an Amicon concentrator (Centricon-10, Amicon Inc., Danvers, MA) to 0.5 ml and passed through a 0.22-μm filter (MILLEX-GV, Millipore Corp., Bedford, MA). The filtered solution was layered on a Superose 12 column (25 ml bed volume, Pharmacia-LKB Biotechnology Inc., Piscataway, NJ) equilibrated with buffer A containing 75 mM NaCl. The column was run at 0.35 ml/min and fractions of 0.5 ml were collected. p66-containing fractions were pooled and stored at −70°C (Pool III). Recovery in Pool III was 0.1–0.2 mg/g of packed cells.

Western Blot Analysis

Proteins were resolved by SDS–PAGE and transferred onto nitrocellulose paper as described (15). The paper was treated with goat whole serum in TBST buffer and reacted with monoclonal antibody (mAb) RT (New England Nuclear) at a 1:1000 dilution for 1 h. After washing the paper with TBST, it was incubated with biotinylated anti-mouse IgG (Cappel Laboratories, Durham, NC) at a 1:500 dilution for 1 h. After washing with TBST, the paper was incubated with an avidin-biotinylated horseradish peroxidase complex (Vectastain ABC Kit, Vector Laboratories, Burlingame, LA) for 1 h. Finally the paper was washed with TBS and color was developed by reaction with 0.3% HRP color development reagent (Bio-Rad Laboratories) and H_2O_2. Protein concentration was determined by the Bradford assay (16). SDS–PAGE of proteins was performed according to Laemmli (17). Protein markers were from Bio-Rad Laboratories.

Reverse Transcriptase Assays

Enzyme activities were determined using a standard reaction mixture (30 μl) containing 50 mM Tris–HCl, pH 7.4, 100 mM KCl, 10 mM $MgCl_2$, 30 μM [α-^{32}P]dTTP, and 1 μM poly(rA)-(dT)$_{16}$ (expressed as 3'-hydroxyl primer termini). Reactions were initiated by addition of cell extracts, incubated at 25°C for 5 min and stopped by addition of 15 μl of 0.5 M EDTA, pH 8.0. Quenched reactions mixtures were spotted onto Whatmann DE-81 filter discs and dried (Whatman, Inc., Clifton, NJ). Unincorporated [α-^{32}P]dTTP was removed by four washes of 0.3 M ammonium formate, pH 8.0, followed by two washes with 95% ethanol and one wash of acetone. The dried filters were counted in 5 ml RPI Bio-safe II.

RESULTS

After cotransfection of pVL941/RT and AcMNPV into Sf9 cells was performed and homologous recombination occurred, recombinant viruses were collected and isolated by plaque purification. Dot blot screening of putative HIV-1 RT recombinant plaques confirmed the presence of the HIV-1 RT cDNA in several of the plaques that were picked for study. A second round of purification was performed using the plaque that gave the strongest signal (H12). RT activity was determined on Sf9 cells after the second round of plaque purification of the plaques from clone H12. The results showed enzymatic activity in the crude cell lysates from the selected recombinant plaques (Fig. 2).

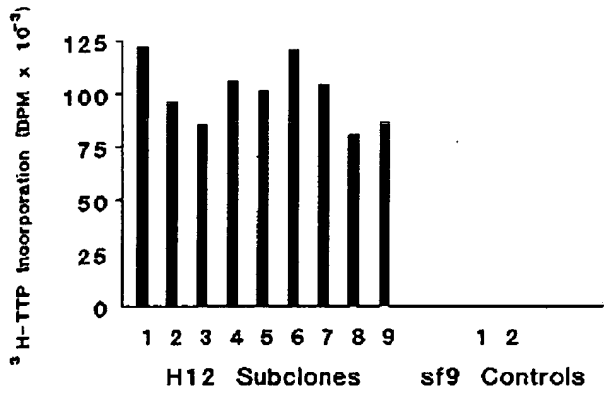


FIG. 2. RT activity expressed during second-round purification of H12 plaques. Results show enzymatic activity in the crude cell lysates of Sf9 cells 48 h postinfection with RT-recombinant baculovirus (H12 subclones), and no activity in noninfected Sf9 cells as control. Each bar represents the RT activity in a lysate of approximately 6×10^5 cells.

Immunoblot analysis of Sf9 cell lysates after infection with recombinant AcMNPV chosen from the second round of plaque purification confirmed the presence of the HIV-1 RT (Fig. 3). Extracts contained an immunoreactive polypeptide of approximately 66 kDa (p66). Mobility of the p66 polypeptide upon SDS–PAGE was similar to that of the p66 HIV-1 RT expressed in *E. coli* (11).

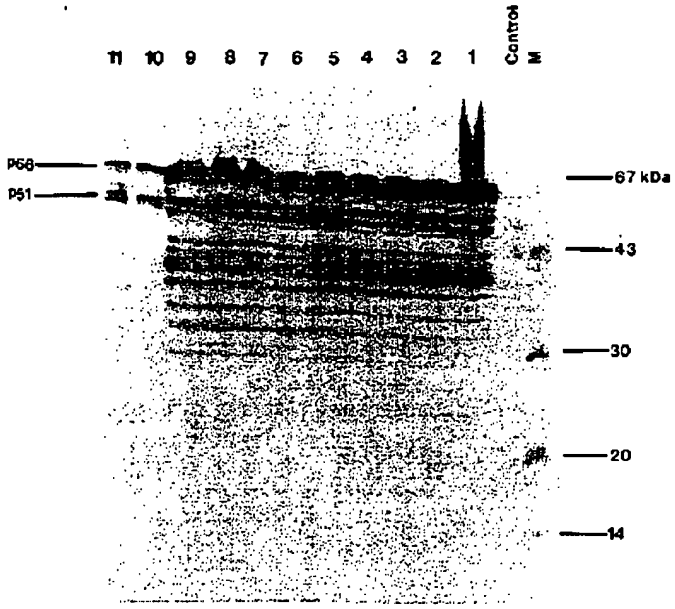


FIG. 3. Immunoblot analysis of Sf9 cell lysates infected with RT-recombinant baculovirus (H12 subclones), using an mAb against RT. Extracts of Sf9-infected cells contained a immunoreactive polypeptide of approximately 66 kDa that corresponds to the RT. Lanes 1–9: Samples of Sf9-infected cells picked after second round plaque purification. Lanes 10 and 11: Partially purified HIV-1 RT expressed in *E. coli*. Control, noninfected Sf9 cells.

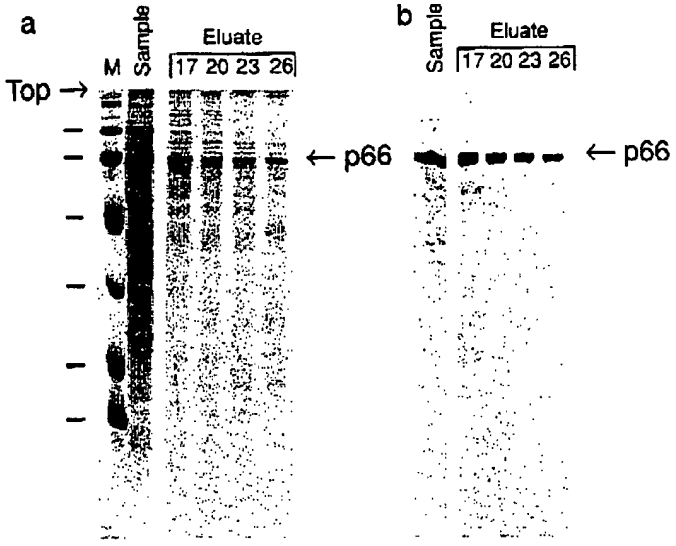


FIG. 4. SDS–PAGE and immunoblot of partial purified HIV-1 RT expressed in Sf9 cells after infection with recombinant baculovirus. The enzyme was recovered in the flowthrough fractions from Mono Q column chromatography (Pool I) and then was chromatographed on a ssDNA–cellulose column. The enzyme bound to the column and was eluted by a sodium chloride gradient. (a) Twenty microliters of the Q Sepharose flowthrough sample (Pool I) and ssDNA–cellulose column eluate (fractions 17–26) was mixed with sample buffer and applied to a 12.5% SDS gel. The baculovirus expressed RT corresponds to a doublet of peptides at approximately 66 kDa. (b) Immunoblot with a 1:100 dilution of mAb to HIV-1 RT confirms the presence of the RT polypeptide. Sample and eluate are the same as described for (a).

RT was purified from the crude extract of baculovirus-infected Sf9 cells following a modification of the procedure used previously for RT expressed in *E. coli* (11). The enzyme was recovered in the flowthrough fractions from Mono Q column chromatography and then was chromatographed on a single-stranded DNA cellulose column. The enzyme bound to the column and was eluted by a sodium chloride gradient. SDS–PAGE analysis of fractions obtained during this partial purification is illustrated in Fig. 4a. The presence of the RT polypeptide was confirmed in fractions using Western blotting with a mAb to RT (Fig. 4b). The RT polypeptide migrated in the gel slightly faster than the bovine serum albumin marker and, therefore, exhibited electrophoretic mobility similar to that of the *E. coli*-expressed RT. In contrast to the *E. coli*-expressed RT, the baculovirus RT frequently corresponded to a doublet of peptides at approximately 66 kDa, rather than a single polypeptide species at this size. The baculovirus-expressed RT was further purified by gel permeation chromatography using a Superose-12 column on an FPLC system. The enzyme emerged from the column in the position expected for a homodimer of the 66-kDa polypeptide(s). RT in the pooled fractions from this column was used for further characterization. Table 1 summarizes protein concentration, total protein, and amount of RT in

TABLE 1

Purification of HIV-1 Reverse Transcriptase from Sf9 Cell Extracts[a]

Fraction	Volume (ml)	Protein concentration (mg/ml)	Total protein (mg)	Reverse transcriptase (mg)
1. Crude extract	100	1	100	3
2. Q Sepharose column (flowthrough)	130	0.35	45.5	2.5
3. ssDNA–cellulose column (Pool)	40	0.13	5.2	1.8
4. Superose 12 column (Pool)	4	0.25	1	1

[a] The experiment was conducted with 10 g of cell pellet. Protein concentrations were measured by the Bradford assay. Amount of RT in each fraction was estimated after SDS–PAGE and Coomassie blue staining. RT migrates in SDS–PAGE just ahead of the 66-kDa marker (bovine serum albumin), and protein estimation was carried out by comparing the intensity of RT staining with that of known amounts of the marker.

each fraction obtained from the purification steps. With respect to p66, this sample was judged to be over 90% pure by SDS–PAGE and Coomassie blue staining (Fig. 5). The most abundant minor contaminating peptide was at 51 kDa. The final fraction of purified RT had a k_{cat} of 0.21/s, compared with 0.33/s measured for homogeneous *E. coli*-derived homodimeric RT. The specific activity of the baculovirus express RT was 103 Units/mg protein, compared to the value obtained for the *E. coli* derived RT of 160 Units/mg protein (Unit = amount of enzyme required to catalyze the incorporation of 1 nmol of dTMP in 1 min at 25°C using poly(rA)-oligo$(dT)_{16}$ as template–primer). The baculovirus-derived enzyme, therefore, has specific activity similar to that of the *E. coli*-derived RT.

DISCUSSION

Both prokaryotic (18) and eukaryotic (13,19) expression systems have been successfully used to generate large quantities of foreign gene products. Expression of the HIV-1 RT has been achieved in bacteria by several laboratories (4–11). Many biologically important proteins, including β-galactosidase (20), c-myc (21), and the human glucocorticoid receptor (hGR) (22) have been expressed in the BEVS. The BEVS has also been demonstrated to be capable of proper post-translational modifications of foreign gene products (23). Post-translational modifications that have been reported to occur in baculovirus-infected insect cells include signal cleavage, proteolytic cleavage, N-glycosylation, O-glycosylation, acylation, amidation, phosphorylation, prenylation, and carboxymethylation (24). It thus appeared that the BEVS would be a proper system that would allow expression of the HIV-1 RT cDNA in an eukaryotic cell, thereby taking advantage of the pathways in these cells that facilitate folding, modification, and assembly of protein products.

After transfection, Sf9 cells were demonstrated to contain the HIV-1 RT cDNA by dot blotting, using as a probe HIV-1 RT cDNA. Western blot analysis indicated the presence of an immunoreactive polypeptide of 66 kDa that corresponded to the predicted size of the transfected cDNA. Enzymatic activity was detected on crude cell lysates of transfected cells. After partial purification (Fig. 4a), the RT expressed in Sf9 cells corresponded to a doublet of approximately 66 kDa, rather than a single polypeptide as it is seen with *E. coli*-expressed RT (11). The origin or basis of this microheterogeneity is unknown, yet a possibility particularly in the BEVS is a post-translational modification of the protein. A system which is designed to express a heterologous gene at a high rate may overwhelm the ability of the cell to modify the protein product, and as a result the post-translational modification of the enzyme may be heterogeneous. This often seems to be the case for the baculovirus system in which the N-glycosylation

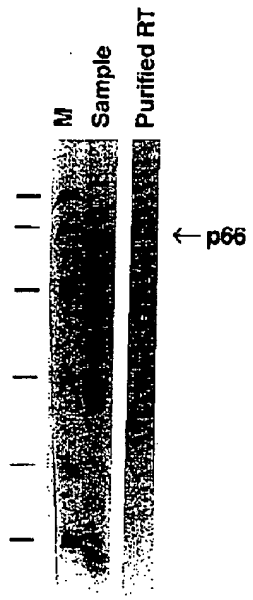


FIG. 5. SDS–PAGE and Coomassie blue stain of the final purified product of the baculovirus expressed RT by gel permeation chromatography using a Superose-12 column in an FPLC system. Sample represents the pool of eluate after ssDNA–cellulose column (Pool II), which was applied on the Superose 12 column. Pool III represents the purified RT. Twenty micrograms of the sample and 4 μg of purified RT were applied to a 12.5% SDS gel. The purified RT emerged from the column in the position expected for a homodimer of the 66-kDa polypeptide. The most abundant minor contaminating peptide was at 51 kDa. M, Molecular mass markers as follows: Phosphorylase b, 97.4 kDa; bovine serum albumin, 66.2 kDa; ovalbumin, 42.7 kDa; carbonic anhydrase, 31 kDa; soybean trypsin inhibitor, 21.5 kDa; lysozyme, 14.4 kDa.

and phosphorylation systems may fail to keep pace with the very high level of expression occurring in the last phase of the infection process (25). Alternatively, after folding into a homodimer, one subunit may be resistant to modification.

Further purification by gel permeation chromatography resulted in the p66 protein. This final product was then characterized for its enzymatic activity, and results were compared to the values obtained with the *E. coli*-derived RT. Our results showed that the baculovirus-derived RT displays characteristics similar to those of the *E. coli*-derived RT. We conclude that RT purified from Sf9 cells will be suitable for structure–function studies of the HIV-1 RT.

REFERENCES

1. Gilbon, E., Mitra, S. W., Goff, S., and Baltimore, D. (1979) A detailed model of reverse transcription and test of crucial aspects. *Cell* **18**, 93–100.
2. Di Marzo Veronese, F., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C., and Sarngadharan, M. G. (1986) Characterization of highly immunogenic p66/pS1 as the reverse transcriptase of HTLV-III/LAV. *Science* **231**, 1289–1291.
3. Lightfoote, M., Coligan, J., Folks, T., Fauci, A. S., Martin, M. A., and Venkatesan, S. (1986) Structural characterization of reverse transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus. *J. Virol.* **60**, 771–775.
4. Hansen, J., Schultze, T., and Moelling, K. (1987) RNase H activity with bacterially expressed reverse transcriptase of human T-cell lymphotropic virus III/lymphadenopathy-associated virus. *J. Biol. Chem.* **262**, 12393–12396.
5. Larder, B., Purifoy, D., Powel, K., and Darby, G. (1987) AIDS virus reverse transcriptase defined by high level expression in *Escherichia coli*. *EMBO J.* **6**, 3133–3137.
6. Hizi, A., McGill, C., and Hughes, S. H. (1988) Expression of soluble, enzymatically active, human immunodeficiency virus reverse transcriptase in *Escherichia coli* and analysis of mutants. *Proc. Natl. Acad. Sci. USA* **85**, 1218–1222.
7. LeGrice, S., Benick, V., and Mous, J. (1987) Expression of biologically active human T-cell lymphotropic virus type III reverse transcriptase in *Bacillus subtilis*. *Gene* **55**, 95–103.
8. Tanese, N., Sodroski, J., Haseltine, W., and Goff, S. P. (1986) Expression of reverse transcriptase activity of human T-lymphotropic virus type III (IITLV-III/LAV) in *Escherichia coli*. *J. Virol.* **59**, 743–745.
9. Farmerie, W. G., Loeb, D. D., Casavant, N. C., Hutchinson, C. A., III, Edgell, M. H., and Swanstrom, R. (1987) Expression and processing of the AIDS virus reverse transcriptase in *Escherichia coli*. *Science* **236**, 305–308.
10. Mous, J., Heimer, E. P., and LeGrice, S. F. (1988) Processing protease reverse transcriptase from human immunodeficiency virus type I polyprotein in *Escherichia coli*. *J. Virol.* **62**, 1433–1436.
11. Becerra, S. P., Kumar, A., Lewis, M. S., Widen, S. G., Karawya, E. M., Abbotts, J., Hughes, S. H., Shiloach, J., and Wilson, S. H. (1991) Protein–protein interactions of HIV-1 reverse transcriptase: Implications of central and C-terminal regions in subunit binding. *Biochemistry* **30**, 11707–11719.
12. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Jr., Pearson, M. L., Lautenberg, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985) Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**, 277–284.
13. Sumers, M. D., and Smith, G. R. (1987) A manual of methods for the baculovirus vectors for insect cell culture procedures. Texas Agricultural Experiment Station, Bulletin No. 155.
14. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) "Molecular Cloning. A Laboratory Manual," pp. 89–96, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
15. Twobin, J., Staehelin, T., and Gordon, J. (1970) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
16. Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–253.
17. Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
18. Dahlman, K., Stromstedt, P. E., Rae, C., Jornval, H., Flock, J. I., Carlstedi-Duke, J., and Gustafsson, J. A. (1989) High level expression in *Escherichia coli* at the DNA-binding domain of the glucocorticoid receptor in a functional form utilizing domain specific cleavage of a fusion protein. *J. Biol. Chem.* **264**, 804–809.
19. Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C. M., and Huang, H. V. (1989) Sindbis virus: An efficient, broad host range vector for gene expression in animal cells. *Science* **243**, 298–301.
20. Pennock, G. D., Shoemaker, C., and Miller, L. K. (1984) Strong and resulted expression of *Escherichia coli* β-galactosidase in insect cell with a baculovirus vector. *Mol. Cell. Biol.* **4**, 399–406.
21. Miyamoto, C., Smith, G. E., Farrell-Towt, J., Chizzonite, R., Summers, M. D., and Ju, G. (1985) Production of human c-*myc* protein in insect cells infected with a baculovirus expression vector. *Mol. Cell. Biol.* **5**, 2860–2865.
22. Srinivasan, G., and Thompson, E. B. (1990) Overexpression of full-length human glucocorticoid receptor in *spodoptera frugiperda* cells using the baculovirus expression vector system. *Mol. Endo.* **4**, 209–216.
23. Fraser, M. J. (1989) Expression of eucaryotic genes in insect cell cultures. *In Vitro Cell. Dev. Biol.* **25**, 225–235.
24. Luckow, V. A. (1991) Cloning and expression of heterologous genes in insect cells with baculovirus vectors, *in* "Recombinant DAN Technology and Applications" (Prokop, A., Ed.), pp. 97–152, McGraw-Hill, New York.
25. O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992) Review of post-translational modification in baculovirus infection, *in* "Baculovirus Expression Vectors," pp. 216–236, Freeman, New York.

Volume 4, Number 4
August 1993

PEXPEJ
ISSN 1046-5928

Protein Expression and Purification

ACADEMIC PRESS, INC. *Harcourt Brace & Company*
San Diego New York Boston London Sydney Tokyo Toronto

Protein Expression and Purification

Volume 4, Number 4, August 1993

1046-5928/93 $5.00

MADE IN THE UNITED STATES OF AMERICA

This journal is printed on acid-free paper.

PROTEIN EXPRESSION AND PURIFICATION

(ISSN 1046-5928)

Published bimonthly by Academic Press, Inc., 6277 Sea Harbor Drive, Orlando, FL 32887-4900

1993: Volume 4. Price: $150.00 U.S.A. and Canada; $177.00 all other countries
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publishers. All correspondence and subscription orders should be addressed to the office of the Publishers at 6277 Sea Harbor Drive, Orlando, FL 32887-4900. Send notices of change of address to the office of the Publishers at least 6 to 8 weeks in advance. Please include both old and new addresses. POSTMASTER: Send changes of address to *Protein Expression and Purification*, 6277 Sea Harbor Drive, Orlando, FL 32887-4900.

EXPRESSION OF ACTIVE HUMAN IMMUNODEFICIENCY VIRUS REVERSE TRANSCRIPTASE IN *SACCHAROMYCES CEREVISIAE*

Philip J. Barr‡, Michael D. Power, Chun Ting Lee-Ng, Helen L. Gibson, and Paul A. Luciw

Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608. ‡Corresponding author.

The reverse transcriptase (RT) domain of the human immunodeficiency virus (HIV) *pol* gene was expressed in yeast. Purification and analysis of the recombinant protein indicated that yeast post-translational processing of RT was similar, if not identical, to that which had previously been observed in HIV infected lymphoid cells. In addition, recombinant RT was enzymatically active and could be inhibited by several anti-retroviral agents. These studies indicate that HIV RT expressed in yeast will be useful for the *in vitro* analysis of further potential inhibitors of this important enzyme.

The etiologic agent responsible for the acquired immunodeficiency syndrome (AIDS) has been identified as a lymphocytopathic retrovirus designated human immunodeficiency virus (HIV). Molecular cloning and sequence analysis has allowed the characterization of three independent isolates previously referred to as lymphadenopathy associated virus (LAV)[1,2], human T-lymphotropic virus type III (HTLV-III)[3,4] and AIDS associated retrovirus (ARV)[5,6]. A criterion for designation of certain viruses as retroviruses is the presence of a viral coded RNA-dependent DNA polymerase (reverse transcriptase) in virions[7,8]. The AIDS retroviral RT has been purified from virions and characterized as highly immunogenic proteins (p66 and p51) detectable on virus Western blots using sera from AIDS patients[9]. Furthermore, RT has been considered to be an excellent target for therapeutic strategies against AIDS virus infection[10,11]. In this study, we show that active RT from cloned HIV-SF2 (a San Francisco isolate of HIV, previously referred to as ARV-2[5,12]) is expressed in genetically engineered *S. cerevisiae* as a similarly processed p66/p51 species and can be purified to near-homogeneity.

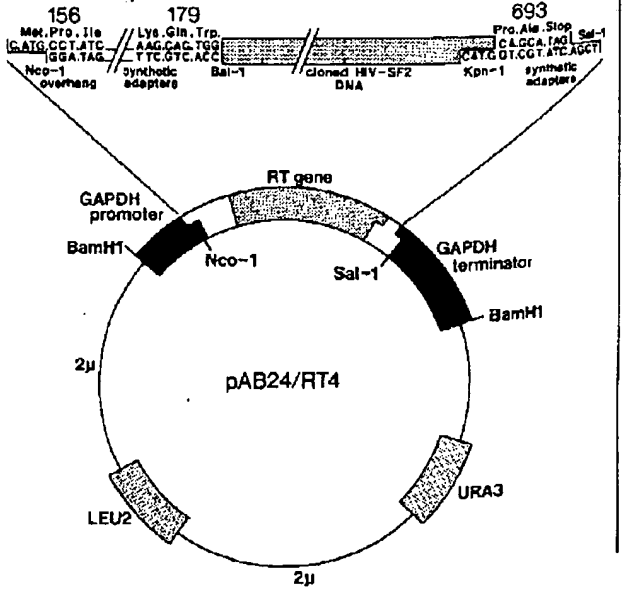
RT is encoded within a domain of the viral *pol* gene[13]. The mature enzyme is derived by proteolytic processing from a large polypeptide precursor whose cleavage is thought to be mediated by a viral protease[13]. This protease, by analogy with other retroviruses, also cleaves the *gag* precursor[13]. RT in virions is present in low quantities[14]. Genetic engineering methodologies offer opportunities to produce large amounts of viral proteins for detailed enzymological and structural studies. DNA fragments encoding portions of the murine leukemia virus (MuLV) RT domain of the pol gene have been cloned into bacterial expression vectors[14,15], and more recently, HIV *pol* gene fusions have been shown to exhibit RT activity when expressed in *E. coli*[16]. We chose to directly express HIV RT in *S. cerevisiae*, since direct expression of eukaryotic proteins in yeast has been shown to be efficient and also to lead to appropriately post-translationally modified proteins[17–19]. Also, RT from cauliflower mosaic virus (CaMV) has been shown to accumulate in an active form in this host[20].

RESULTS AND DISCUSSION

For direct expression of the RT domain in yeast, we estimated the *N*- and *C*-termini of the mature protein by drawing on homology comparisons with the amino acid sequences of *pol* gene products of other retroviruses. Precise amino acid choices for termini were based on the target specificities of retroviral proteases, including the AIDS virus protease, from known *gag* subunit sequences. Accordingly, the Phe.Pro at amino acid positions 155 and 156 of the ARV-2 *pol* open reading frame[6] and the Val.Pro at 163/164 were selected as likely *N*-termini. A likely *C* terminal processing site was estimated at Val.Pro 691/692. For construction of yeast expression vectors, we used the naturally occurring restriction sites Bal-1 at Trp 179 and the Kpn-1 site at Trp 690 together with synthetic oligonucleotide adapters to fuse the desired coding regions to yeast control sequences (Fig. 1). Since the *N*-terminal sequence of HIV RT has been reported[6], and is indeed Pro 156, we will describe only expression results from constructions at this processing site.

For high level constitutive expression of HIV-SF2 reverse transcriptase in yeast we constructed a vector in which the RT gene was flanked by promoter and terminator sequences of the yeast glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene[17] (Fig. 1). Yeast cells transformed with this autonomously replicating yeast plasmid, designated pAB24/RT4, were analyzed by poly-

FIGURE 1 Yeast plasmid construction for expressing active HIV RT. See Experimental Protocol for details.



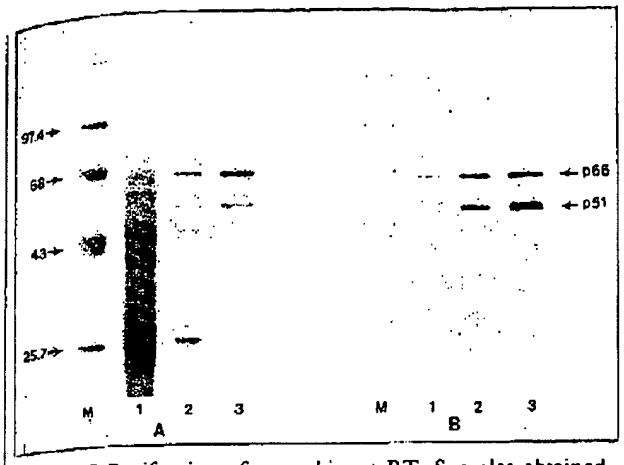


FIGURE 2 Purification of recombinant RT. Samples obtained after each purification step (see Experimental Protocol) were analyzed by SDS-PAGE using coomassie blue staining (Panel A) or Western blotting with a single human AIDS serum (Panel B). M: protein molecular weight markers; Lane 1: proteins in the 0–30% NH_4SO_4 pellet; Lanes 2 and 3: peak RT fractions after cellulose phosphate and DNA-cellulose chromatography respectively.

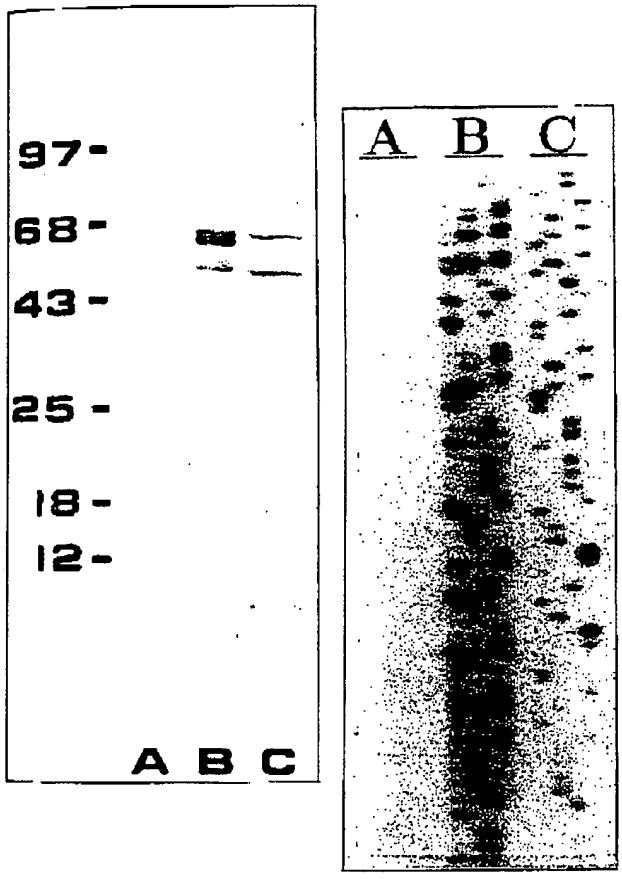


FIGURE 3 Immunoblot of viral and recombinant RT. Infected cell lysate (A), yeast recombinant RT (B) and soft pelleted virus (C) were run on a 10–20% gradient SDS gel prior to immunoblotting. Infected cell lysate and soft pelleted virus were prepared as previously described[21]. Recombinant RT was prepared as described in Experimental Protocol. Rabbit antibody to prt/RT was as described in the text.

FIGURE 4 Sequencing of M13 template using control yeast extract (lane A), crude lysate from yeast cells harboring plasmid pAB24/RT (lane B), DNA polymerase, Klenow fragment (lane C) using standard conditions[32].

acrylamide gel electrophoresis and immunoblotting techniques using an AIDS patient's serum (Fig. 2). Reactivity to HIV RT purified from virions was previously demonstrated to be present in at least 80% of HIV antibody positive sera tested[9]. Such immunogenicity of a retroviral reverse transcriptase has not been reported previously in a natural retroviral infection. Preliminary immunoblot analysis of recombinant RT also indicates a similar high degree of reactivity with AIDS sera. Of 20 sera tested, 19 scored positive (data not shown). Availability of recombinant HIV RT will thus allow serological correlations of immunological phenomena associated with AIDS retrovirus infection and replication, including analysis of disease stage-specific antibody responses to this individual antigen. We have previously demonstrated such utility for other domains of the AIDS virus produced in recombinant microorganisms[21]. A full serological analysis of RT derived from yeast is currently in progress.

RT purified from virus particles consists of approximately equimolar amounts of two species designated p66 and p51[9]. These proteins have identical *N*-termini, and presumably differ at their C-termini. If p66 and p51 are present as a heterodimeric complex, the association does not involve disulfide linkage[9]. RT activity for either p66 or p51 alone has not been reported. Our construction encodes a protein with a calculated molecular weight of 62.5 kD. Figure 2, Panel A, shows a polyacrylamide gel analysis of yeast cells expressing HIV RT, and its three step purification to apparent homogeneity. The apparent gel mobility of the major component as approximately 66kD indicates a close approximation to the natural p66 alone. It was also noted that during purification, processing of this 66kD protein occurred giving a second major species with an estimated molecular weight of 51 ± 1.5kD. This processing is presumably due to a yeast protease since the region thought to encode the HIV protease was not included in this expression construction. Previously, HIV *gag-pol* fusions expressed in yeast were shown to be processed when this protease region was included[18]. Recently, a potential C-terminus for the larger RT species was defined by *N*-terminal sequencing of the integrase protein immediately adjacent to RT in the *gag-pol* precursor[22].

That the p66 and p51 species produced in yeast had identical *N*-termini (Pro.Ile.Ser.Pro.Ile etc.) was confirmed by 15 cycles of gas phase sequence analysis of the purified proteins (Fig. 2, Panel A, lane 3). To further confirm that p51 was derived from p66 by C-terminal processing, we used rabbit antibody prepared against a polypeptide representing amino acids 79 to 192 of the *pol* gene. This large peptide was produced by secretion from yeast using the α-factor mating pheromone leader sequence[23,24] and represents 77 amino acids of the putative protease region and 37 N-terminal amino acids of RT[9]. This antibody (prt/RT antibody) readily detected yeast derived p66 and p51 on a Western blot (Figure 3, lane B) confirming the presence of these *N*-terminal amino acids in p51. Thus, yeast may mimic the natural maturation processes for HIV RTs, giving rise to both p66 and p51. The sequence analysis also showed that the *N*-terminal methionine derived from the synthetic initiation codon was removed *in vivo*. Interestingly, the C-terminal processing observed in yeast may indicate that HIV protease is not necessarily required for this processing event *in vivo*. Confirmation of the identity of the purified proteins as HIV related is shown in Figure 2, Panel B. Western blot analysis using an AIDS patient's serum showed that processing occurred during purification and also resolved the 51kD protein into two bands. No HIV-related polypeptides were detected with this serum in control yeast

TABLE 1 Purification of recombinant HIV RT.

Fractionation Step	Total Protein (mg)	Total Units	Specific Activity (U/mg)
Crude Lysate	1,575	28,586	18.2
Phosphocellulose chromatography	1.995	21,433	10,743
Single-stranded DNA-chromatography	0.356	12,341	34,666

Assays were performed as previously described[9] except that dTTP was at 136.5 μM. 1 unit of the enzyme catalyzes the incorporation of 1 nmole of dTMP in 60 min. at 37°C.

TABLE 2 Inhibition of HIV RT.

Inhibitor	Concentration	cpm Incorporated/h	% Inhibition
AZTTP	0.01 μM	39,237	53.5
	0.1 μM	8,550	90.7
	1.0 μM	1,816	98.9
HPA23	0.005 μM	77,568	7.0
	0.05 μM	6,674	93.0
	0.5 μM	858	100.0‡
PFA	0.05 μM	53,036	36.8
	0.5 μM	16,483	81.1
	5.0 μM	1,562	99.2
No inhibitor		83,414	—
Background		934	—

‡HPA-23 at concentrations of 0.5 μm gave slightly lower incorporation than baseline values.
Purified recombinant RT was assayed as described previously[9] except that the dTTP concentration was 0.17 μM, in the presence of the above RT inhibitors. Enzyme and label were preincubated with the inhibitor prior to the addition of primer/template substrates. The molecular weight of HPA23 was calculated from the monomeric $(Sb_2W_5O_{20})(NH_4)_4.8H_2O$, formula weight of 1698.8[27]. Counts per minute (cpm) incorporated per hour were measured at 60 minutes and are the average of three reactions. PAA was not inhibitory at concentrations as high as 5.0 μM. Ansamycin failed to inhibit recombinant RT at 590 μM.

samples (data not shown). We also compared, by immunoblotting, a preparation of recombinant RT from yeast with lysates from HIV infected HUT-78 cells and purified HIV (Fig. 3). In these samples, several species are observed in both the yeast and virus preparations. Yeast p66 and p51 appear to be identical in gel electrophoretic mobility to the virion derived proteins (lanes B and C). The infected cell lysate contains a polypeptide at *ca.* 66kD (lane A) but slightly smaller than the p66 of purified virus. Strikingly, infected cells do not contain a p51 species detectable by this procedure. The simplest explanation for our observation is that p66 observed in purified virus may arise from cleavage at the reported *N*-termini of RT[9] and the integrase[22], followed by *C*-terminal processing to produce the slightly smaller 66kD protein observed in infected cells and p51 in virions. Of course, we cannot rule out other post-translational modification mechanisms to account for the various forms of RT observed. Interestingly, the yeast RT in this preparation (Fig. 3, lane B) contains species comigrating with all the above bands. Lastly, we observed an immunoreactive species in the virus preparation at around 10kD using the prt/RT antibody. Although we cannot formally eliminate the possibility of this being a small RT related peptide, it is likely that this represents the product of the HIV protease coding region *(prt)*. Experiments are currently in progress to further analyze this important aspect of the HIV virion structure.

RT activity. Final analysis of the recombinant RT was by enzymatic assay of both crude yeast lysates and the purified enzyme. Using the assay conditions described for RT isolated from virions[9], the relative activity of yeast derived RT, from crude lysates, was assayed using various primer template combinations. In all cases, reactions were linear for greater than 30 minutes. Thus, relative to $(dT)_{\sim 15} \cdot (rA)_n$ (100%) the enzyme activities with $(dT)_{\sim 15} \cdot (dA)_n$ was 4.3%, $(dG)_{\sim 15} \cdot (rC)_n$, 71.1% and $(dG)_{\sim 15} \cdot (rC^m)_n$, 2.4%. Yeast extracts from cells containing control plasmids gave background levels of incorporation, excluding the possibility of host encoded RT activity. RT activity was measured throughout purification (Table 1) and a final specific activity of around 3.5×10^4 U/mg at 37°C obtained. This value is somewhat higher than that observed previously for cauliflower mosaic virus RT expressed in yeast (*ca.* 4×10^3 U/mg at 20°C)[20].

In order to qualitatively determine polymerizing fidelity of recombinant HIV RT, we compared the use of crude lysates from yeast expressing HIV RT in an M13 dideoxy sequencing reaction with the conventional procedure using bacterial DNA polymerase (Fig. 4). The resulting sequence ladder demonstrated the polymerization ability of recombinant RT and also showed the inhibition of chain length extension by dideoxynucleotides *in vitro*. The appearance of several additional "ghost" bands in the RT derived sequence (lane B) may reflect an increased error frequency for HIV RT as compared with Klenow (lane C). A detailed analysis of this observation may contribute to our understanding of the high sequence variability of this virus.

Inhibition of HIV RT. 2',3'-Dideoxynucleosides have recently been shown to inhibit replication of the AIDS virus *in vitro*[11]. This inhibition was extended to the nucleotide analog 3'-azido-3'-deoxythymidine-5'-triphosphate (AZTTP), the probable active metabolite of 3'-azido-3'-deoxythymidine (AZT), a RT inhibitor with demonstrated antiviral activity towards the AIDS virus both in *in vitro* and clinical settings[9,25] (Table 2). We also studied the effects of other drugs currently in clinical trials against AIDS on the purified recombinant RT. As shown in Table 2, using purified RT in a preliminary assay system, several of the inhibitors tested; AZTTP, phosphonoformate (PFA)[26], and HPA23[27] were very inhibitory at clinically relevant concentrations. Phosphonoacetate (PAA) and ansamycin[28] were not inhibitory at these or higher concentrations (Table 2, legend). Surprisingly, ansamycin had previously been shown to inhibit RT from other retroviruses[29]. This raises the possibility of alternate modes of inhibition of HIV by ansamycin *in vivo*. Rifampicin derivatives have previously been shown to inhibit vaccinia virus replication by disruption of virion morphogenesis[30]. Binding of ansamycin to RT might similarly affect HIV virion assembly, and consequent cytopathogenicity.

Production of large quantities of this recombinant protein in yeast will allow a full and more detailed enzymological analysis of the kinetics and inhibition of HIV RT *in vitro* as compared with mammalian cell polymerases, and may also help in elucidating the relative functions, if any, of each individual species (p66 and p51) of HIV RT. Thus, these results show the utility of HIV RT produced in recombinant yeast cells as a reagent for both immunological studies and antiviral chemotherapeutic approaches towards inhibiting AIDS retrovirus replication.

EXPERIMENTAL PROTOCOL

Yeast plasmid construction for the expression of active HIV RT. A Bal-1/Kpn-1 fragment of cloned proviral ARV-2 DNA[6] encompassing amino acids Pro 180 and Trp 690 of the *pol* open reading frame[6] was excised from a pUC19 plasmid containing this region of the cloned virus[6]. This fragment was extended,

using synthetic oligonucleotides, to include Pro 156 at the *N*-terminus and Ala 693 at the C-terminus. The synthetic DNA also provided a methionine initiation codon. Oligonucleotides were synthesized by the phosphoramidite method using Applied Biosystems 380A DNA synthesizers. Linkers and cloned DNA were ligated by standard techniques[31] into the vector pPGAP[17], a plasmid containing yeast glyceraldehyde-3-phosphate (GAPDH) promoter and terminator sequences flanking unique Nco-1 and Sal-1 cloning sites. The resulting promoter-gene-terminator expression cassette was excised using BamH1, and cloned into BamH1 digested and alkaline phosphatase treated pAB24. This autonomously replicating yeast plasmid contains LEU2 and URA3 markers for selection and 2 micron DNA sequences for autonomous replication in yeast and will be described in detail elsewhere (A. J. Brake, unpublished). The resulting vector, designated pAB24/RT4 was used to transform yeast strain AB110[17] and leucine prototrophs grown in leucine deficient media[17] to 3ml, followed by growth in YEPD[17] to the 1 litre level.

Purification of recombinant RT. Cells from a 1 liter culture were pelleted by centrifugation at 2,500 rpm for 10 minutes. The cell pellet was resuspended in 300 ml of 50mM Tris-HCl, pH 7.5, 14mM β-mercaptoethanol, 1.2M sorbitol, and 200 micrograms/ml Zymolyase. Spheroplast formation, monitored by light microscopy, was allowed to proceed for 90 minutes at 30°C. After a low speed centrifugation, the pelleted spheroplasts were lysed in a buffer containing 50mM Tris-HCl, pH 7.5, 0.1% Triton X-100, and 1mM DTT at room temperature. The yeast lysate was clarified by centrifugation at 20,000 rpm for 2 hours and the supernatant was fractionated by step-wise NH_4SO_4 precipitation. Greater than 90% of reverse transcriptase activity was in the 0–30% NH_4SO_4 insoluble fraction. This NH_4SO_4 pellet was resuspended in 20ml of reverse transcriptase buffer (RTB; 50mM Tris-HCl, pH 7.5, 2mM β-mercaptoethanol, 0.2mM EDTA, 0.1% Triton X-100, 20% vol/vol glycerol) containing 50mM KCl. An Amicon pressure filtration device was used for desalting. The extract was then applied to a cellulose phosphate column (Sigma C-2383) (2.5 cm × 30 cm) pre-equilibrated in RTB containing 50mM KCl. The column was washed with 100ml of the same buffer. A linear gradient of 50 to 800 mM KCl in RTB was used for elution. Individual fractions were monitored for RT activity. The peak RT fractions were between 150 and 225 mM KCl. These peak fractions (about 15ml total) were pooled and desalted by Amicon filtration using RTB containing 50mM KCl. This material was then applied to a single stranded DNA cellulose column (Sigma D-8273) (1.0 cm × 10 cm) pre-equilibrated in RTB at 50mM KCl. The column was washed with 30 ml of this buffer and eluted with a linear gradient of RTB from 50 to 800 mM KCl. Fractions were monitored for RT activity and peak fractions were pooled.

Acknowledgments

We thank T. Krenitsky and R. L. Miller of Burroughs-Wellcome, N. Carolina, J. C. Chermann and F. Barré-Sinoussi, Pasteur Institute, Paris and Adria Laboratories, Columbus, Ohio for generously providing the RT inhibitors AZTTP, HPA23 and ansamycin respectively and E. T. Miller for preparation of virus and infected cell lysate. We also thank D. V. Santi, J. A. Levy and H. E. Varmus (UCSF) for helpful discussions, F. Masiarz for protein sequencing and our many friends and colleagues at Chiron Corporation for their support and encouragement. We also acknowledge Toni Jones and Peter Anderson for preparation of the manuscript.

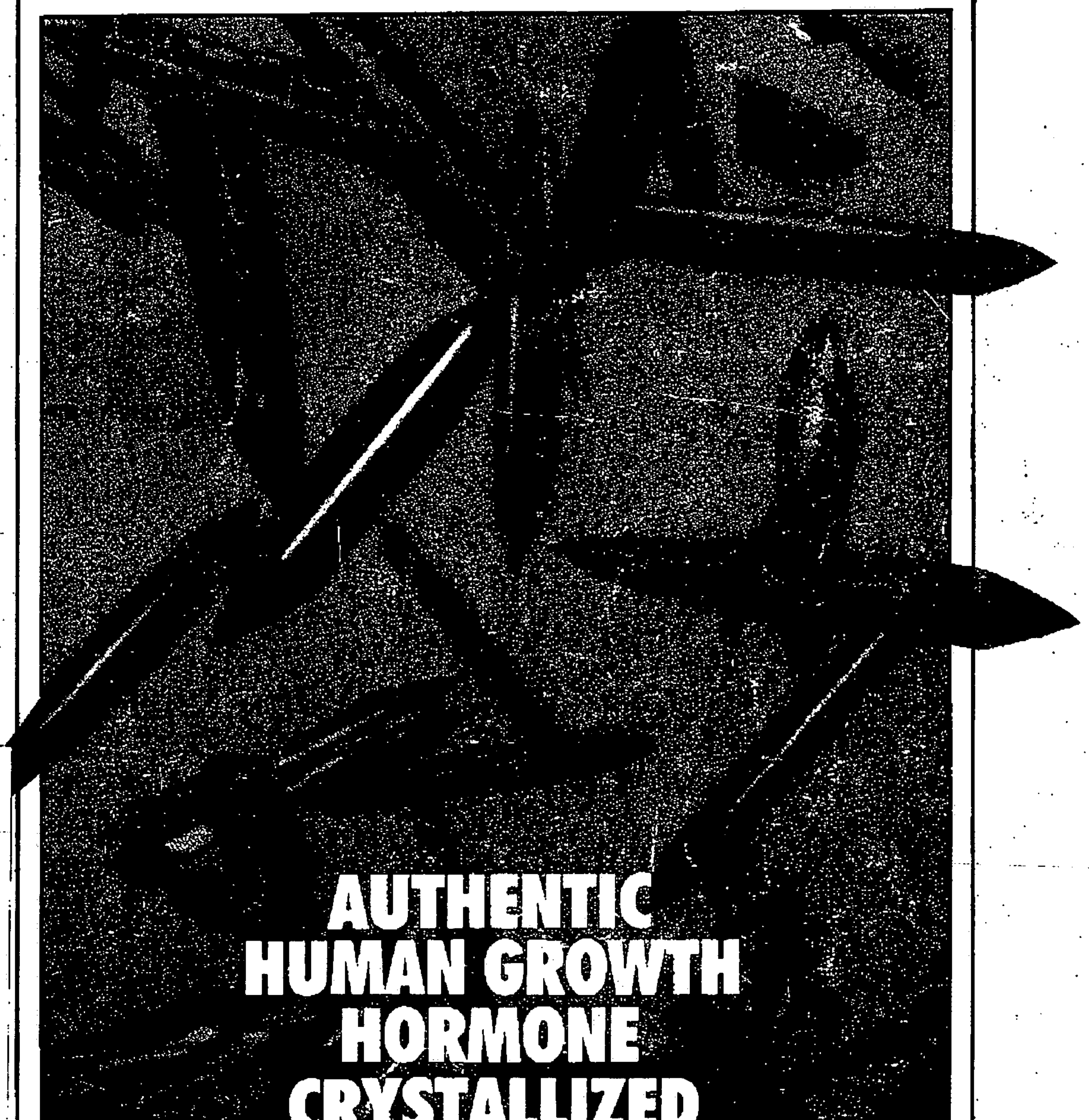
Received 9 December 1986; accepted 18 February 1987.

References

1. Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., and Axler-Blin, L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for Acquired Immune Deficiency Syndrome (AIDS). Science **220**:868–871.
2. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., and Alizon, M. 1985. Nucleotide Sequence of the AIDS Virus, LAV. Cell **40**:9–17.
3. Popovic, M., Sarngadharan, M. G., Read, E., and Gallo, R. C. 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from Patients with AIDS and Pre-AIDS. Science **224**:497–504.
4. Ratner, L., et al. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature **313**:277–284.
5. Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. M., and Oshiro, L. S. 1984. San Francisco patients with AIDS. Science **225**:840–842.
6. Sanchez-Pescador, R., et al. 1985. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). Science **227**:484-492.
7. Baltimore, D. 1970. RNA-dependent DNA polymerase in virions of RNA tumor viruses. Nature **226**:1209–1211.
8. Temin, H. M. and Mitzutani, S. 1970. RNA-directed DNA polymerase in virions of Rous sarcoma virus. Nature **226**:1211–1213.
9. Veronese, F. M., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C., and Sarngadharan, M. G. 1986. Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. Science **231**:1289–1291.
10. Mitsuya, H., Weinhold, K. J., Furman, P. A., St. Clair, M. H., Nusinoff, L., et al. 1985. 3'-Azido-3'-deoxythymidine (BW A509U): An antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus *in vitro*. Proc. Natl. Acad. Sci. USA. **82**:7096–7100.
11. Mitsuya, H. and Broder, S. 1986. Inhibition of the *in vitro* infectivity and cytopathic effect of human T-lymphotrophic virus type III/lymphadenopathy-associated virus (HTLV III/LAV) by 2',3'-dideoxynucleosides. Proc. Natl. Acad. Sci. USA **83**:1911–1915.
12. Luciw, P. A., Potter, S. J., Steimer, K., Dina, D., and Levy, J. A. 1984. Molecular cloning of AIDS-associated retrovirus. Nature **312**:760–763.
13. Weiss, R., Teich, N., Varmus, H., and Coffin J. 1982. RNA Tumor Viruses. Cold Spring Harbor Laboratory, New York.
14. Tanese, N., Roth, M., and Goff, S. P. 1985. Expression of enzymatically active reverse transcriptase in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82**:4944–4948.
15. Kotewicz, M. L., D'Alessio, M. D., Driftmier, K. M., Blodgett, K. P., and Gerard, G. F. 1985. Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *Escherichia coli*. Gene **35**:249–258.
16. Tanese, N., Sodroski, J., Haseltine, W. A., and Goff, S. P. 1986. Expression of reverse transcriptase activity of human T-lymphotropic virus Type III (HTLV III/LAV) in *Escherichia coli*. J. Virol. **59**:743–745.
17. Travis, J., Owen, M., George, P., Carrell, R., Rosenberg, S., Hallewell, R. A., and Barr, P. J. 1985. Isolation and properties of recombinant DNA produced variants of human α_1-proteinase inhibitor. J. Biol. Chem. **260**:4384–4389.
18. Kramer, R. A., Schaber, M. D., Skalka, A. M., Ganguly, K., Wong-Staal, F., and Reddy, E. P. 1986. HTLV-III *gag* protein is processed in yeast cells by the virus *pol*-protease. Science **231**:1580–1584.
19. Hallewell, R. A., Mills, R., Tekamp-Olson, P., Blacher, R., Rosenberg, S., Otting, F., Masiarz, F. R., and Scandella, C. S. 1987. Human Cu, Zn superoxide dismutase produced in yeast: N-terminal acetylation and resistance to copper and zinc. Bio/Technology **5**:363–366.
20. Takatsuji, H., Hirochika, H., Fukushi, T., and Ikeda, J. E. 1986. Expression of cauliflower mosaic virus reverse transcriptase in yeast. Nature **319**:240–243.
21. Steimer, K. S., Puma, J. P., Power, M. D., Powers, M. A., George-Nascimento, C., et al. 1986. Differential antibody responses of individuals infected with AIDS-Associated retroviruses surveyed using the viral core antigen p25*gag* expressed in bacteria. Virology **150**:283–290.
22. Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A., and Venkatesan, S. 1986. Structural characterization of reverse transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus. J. Virol. **60**:771–775.
23. Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., et al. 1984. α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **81**:4642–4646.
24. Barr, P. J., Parkes, D., Sabin, E. A., Power, M. D., Gibson, H. L., et al. 1986. Expression of functional domains of the AIDS associated retrovirus (ARV) in recombinant microorganisms. *In:* U.C.L.A. Symposium on Molecular and Cellular Biology, Vol. 43. Gallo, R. C., Haseltine, W., Klein, G., and Zur Hausen, H. (eds). Alan R. Liss, Inc. New York.
25. Yarchoan, R., et al. 1986. Administration of 3'-azido-3'-deoxythymidine, an inhibitor of HTLV-III/LAV replication, to patients with AIDS or AIDS-related complex. Lancet (i):575–580.
26. Sandstrom, E. G., Kaplan, J. C., Byington, R. E., and Hirsch, M. S. 1985. Inhibition of human T cell lymphotropic virus type III *in vitro* by phosphonoformate. Lancet (i):1480–1482.
27. Chermann, J.-C., Sinoussi, F. C., and Jasmin, C. 1975. Inhibition of RNA-dependent DNA polymerase of murine oncornaviruses by ammonium-3-tungsto-2-antimoniate. Biochem. Biophys. Res. Commun. **65**:1229–1236.
28. Anand, R., Moore, J., Feorino, P., Curran, J., and Srinivasan, A. 1986. Rifabutine Inhibits HTLV-III. Lancet (i):97–98.
29. Gurgo, C., Ray, R., and Green, M. 1972. Rifamycin derivatives strongly inhibiting RNA→DNA polymerase (reverse transcriptase) of murine sarcoma viruses. J. Natl. Cancer Inst. **49**:61–79.
30. Grimley, P. M., Rosenblum, E. N., Mims, S. J., and Moss, B. J. 1970. Interruption by rifampin of an early stage in vaccinia virus morphogenesis: Accumulation of membranes which are precursors of virus envelopes. Virology **6**:519–524.
31. Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. Molecular Cloning, a Laboratory Manual. Cold Spring Harbor, New York.
32. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.

BIO/TECHNOLOGY

THE INTERNATIONAL MONTHLY FOR INDUSTRIAL BIOLOGY



THE LABORATORY ROBOT

BIO/TECHNOLOGY
MAY 1987
VOLUME 5
NUMBER 5

THE FIRST WORD

CHRONICLE

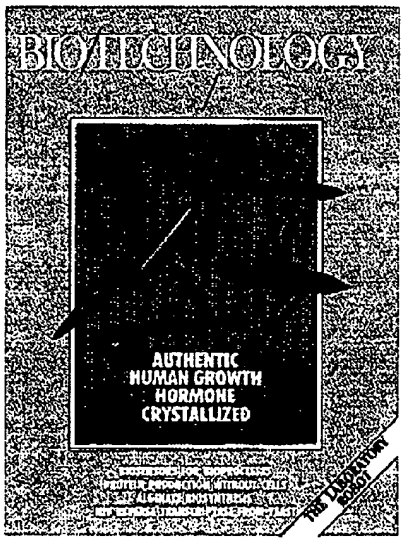
DATELINE

ARTICLES

LAB ROBOTICS

REVIEW

RESEARCH PAPERS



THE COVER photo, courtesy Noel D. Jones et al., Lilly Research Laboratories (Indianapolis, IN), shows large single crystals of natural-sequence recombinant human growth hormone. The needle-like trigonal prisms with pointed ends are suitable for x-ray crystallographic analysis. See page 499.

MATERIALS/METHODS

READER SERVICE

CORRESPONDENCE

CLASSIFIED

THE LAST WORD

BIO/TECHNOLOGY (ISSN 0733-222X) is published monthly by Nature Publishing Co., headquartered at 65 Bleecker, New York, N.Y. 10012-2467, which is owned by Nature America, Inc., a subsidiary of Macmillan Magazines of London.
North American Advertising: BIO/TECHNOLOGY; 65 Bleecker St., New York, New York 10012-2467 (212) 477-9600. **European Advertising:** 4 Little Essex St., WC2R 3LF. Telephone: (01) 836 6633. Telex: 262024. **New Subscriptions, renewals, changes of address, back issues, and all customer service questions** should be addressed to: BIO/TECHNOLOGY, Customer Relations, P.O. Box 1543, Neptune, N.J. 07754-1543. Call toll free: 800-524-0328 **Annual Subscription Rates:** U.S./Canada: US$59; UK: £70; Australia: A$130; Rest of World: US$112. JAPAN: write Japan Publications Trading Company Ltd. 2-1 Sarugaku-cho 1-chome, Chiyoda-ku, Tokyo, Japan. Tel: (03) 292-3755. **Back Issues:** Available at the following rates: U.S./Canada: $9.00; Rest of World: Surface U.S. $11.50; Rest of World: Air Mail US $15.00.
Reader Inquiry Service: BIO/TECHNOLOGY, Nature America, Inc., Reader Service Department, P.O. Box 2679, Clinton, Iowa 52735. **Reprints:** inquiries should be addressed to the editorial office in New York.
POSTMASTER: send address changes to BIO/TECHNOLOGY Subscription Dept., P.O. Box 1543, Neptune, N.J. 07754-1543. Second class postage paid at New York, NY 10012-2467 and additional mailing offices.
Executive Officers of Nature America, Inc.: Nicholas Byam Shaw, Chairman of the Board; Gary M. Reksrad, President; and Edward Valis, Secretary-Treasurer.

PROTEIN EXPRESSION AND PURIFICATION 2, 95–107 (1991)

REVIEW

Fusion Tails for the Recovery and Purification of Recombinant Proteins

Clark F. Ford,* Ilari Suominen,† and Charles E. Glatz‡

*Department of Food Science and Human Nutrition and ‡Department of Chemical Engineering, Iowa State University, Ames, Iowa 50011; and †Department of Biochemistry, University of Turku, Turku, Finland

Several fusion tail systems have been developed to promote efficient recovery and purification of recombinant proteins from crude cell extracts or culture media. In these systems, a target protein is genetically engineered to contain a C- or N-terminal polypeptide tail, which provides the biochemical basis for specificity in recovery and purification. Tails with a variety of characteristics have been used: (1) entire enzymes with affinity for immobilized substrates or inhibitors; (2) peptide-binding proteins with affinity to immunoglobulin G or albumin; (3) carbohydrate-binding proteins or domains; (4) a biotin-binding domain for *in vivo* biotination promoting affinity of the fusion protein to avidin or streptavidin; (5) antigenic epitopes with affinity to immobilized monoclonal antibodies; (6) charged amino acids for use in charge-based recovery methods; (7) poly(His) residues for recovery by immobilized metal affinity chromatography; and (8) other poly(amino acid)s, with binding specificities based on properties of the amino acid side chain. Fusion tails are useful at the lab scale and have potential for enhancing recovery using economical recovery methods that are easily scaled up for industrial downstream processing. Fusion tails can be used to promote secretion of target proteins and can also provide useful assay tags based on enzymatic activity or antibody binding. Many fusion tails do not interfere with the biological activity of the target protein and in some cases have been shown to stabilize it. Nevertheless, for the purification of authentic proteins a site for specific cleavage is often included, allowing removal of the tail after recovery. © 1991 Academic Press, Inc.

Biospecific and other biochemical interactions provide powerful tools for protein recovery and purification, sometimes allowing a specific target protein at a very low initial concentration to be purified in a single step from a complex mixture of proteins. The power of individual biospecific and biochemical interactions can now be brought to bear on the recovery and purification of a wide spectrum of target proteins by genetically engineering the target protein to contain a specific affinity or interactive fusion tail. This topic has recently been the subject of several excellent reviews (1–5). Tails with a variety of sizes and specificities have been designed for fusion to virtually any target protein that can be cloned and expressed in a microbial host. In addition to facilitating recovery, fusion tails have also been used to enhance the stability of small heterologous target proteins (such as peptide hormones) in *Escherichia coli*, preventing their proteolytic degradation, and to provide easily assayable "tags" to monitor the presence of a target protein during purification.

Fusion tails can be genetically engineered either by addition of a restriction fragment encoding the tail to the target protein gene or by inserting the target protein gene into a vector that contains the tail (Fig. 1). Many fusion tail systems have been constructed in specific expression vectors that contain convenient cloning sites for the insertion of the target protein gene in one or more reading frames. Fusion tails have been added to either the N- or the C-terminus of proteins and, in some applications, to both termini. Fusion tails have been designed to facilitate recovery of intracellular or extracellular proteins from a variety of cell types. Specific proteolytic cleavage sites can be genetically engineered into fusion tail systems for removal of the tail from the target protein after purification. In many cases, the tail can then be separated from the target protein using the same interaction or affinity system that was used for initial recovery.

Applications of fusion tail technology range from basic research to industrial production. On a lab scale, fusion tail recovery systems are powerful and elegant

1046-5928/91 $3.00

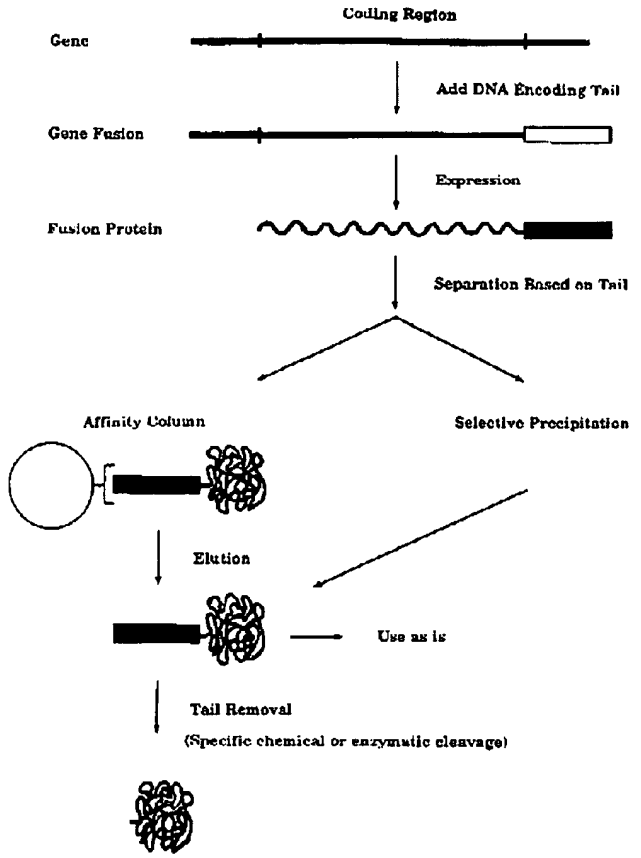


FIG. 1. Construction and use of fusion tails. The tail is shown at the C-terminus in this diagram but many tails can be engineered at the N-terminus of the target protein.

tools for one-step recovery and purification of recombinant proteins or identification of proteins encoded by cloned cDNAs. On an industrial scale, fusion tail technology can be used in the recovery and purification of both higher-cost pharmaceuticals and lower- to medium-cost enzymes.

In this review we will briefly introduce some of the separation systems available for recovery and purification of fusion proteins and will address some of the issues that need to be considered when choosing a fusion tail system. We will not address the topic of protein fusions that are designed primarily to enhance heterologous expression (or stabilization) or where interactions with the fusion tail are not used for purification or recovery (see (1,6,7)).

TAIL TYPES

Tails specifying a variety of different biospecific and biochemical interactions have been utilized as the basis of fusion protein recovery systems (Table 1). These include (1) entire enzymes with affinity for substrates or inhibitors; (2) peptide-binding proteins; (3) carbohydrate-binding proteins or domains; (4) biotin-binding domains; (5) antigenic epitopes with affinity to specific immobilized monoclonal antibodies; (6) charged amino acids for use in charge-based recovery methods; (7) poly(His) residues for binding to metal chelates and recovery by immobilized metal affinity chromatography; and (8) other poly(amino acid)s with various binding specificities.

1. Enzymes

β-Galactosidase (βGal) from *E. coli* (a tetramer with monomer size 116 kDa) has been extensively used as a fusion tail for the recovery of proteins expressed intracellularly in *E. coli* (8–13). Vectors allowing high-level regulated expression are available for construction of fusion proteins at either the N-terminus or the C-terminus of βGal (14–17). One of the most versatile βGal vectors, pMSgt11 (18), is compatible with the λgt11 expression vector, allowing in-frame cloning of λ inserts at an *Eco*RI site in the plasmid. Fusion to βGal has been very effective in protecting heterologous peptides from proteolytic degradation (19–22). The fusion proteins can be insoluble, however, necessitating recovery from inclusion bodies by use of strong denaturing agents. Soluble βGal fusion proteins, on the other hand, can be purified by affinity to substrate analogs (9,23) or by immunoaffinity (24–26). Site-specific proteolysis by either collagenase (8,16,26,27) or blood factor X_a (17,18) has been used to cleave the target protein from βGal.

Glutathione-*S*-transferase (GST) from *Schistosoma japonicum* (26 kDa) is another enzyme that has been utilized as a fusion tail for the isolation of a large number of intracellular target proteins expressed in *E. coli* (28). Although the majority of GST fusions studied have been soluble, some fusions may be insoluble. Soluble GST fusions can be purified on immobilized glutathione and eluted with reduced glutathione under nondenaturing conditions. Insoluble GST fusion proteins can also be affinity-purified if they are first solubilized in low concentrations of ionic or nonionic detergents. Convenient vectors are available for high-level regulated expression of GST fusion proteins from a *tac* promoter. The vectors encode protease cleavage sites for thrombin or blood coagulation factor X_a at the C terminus of GST for removal of the tail after purification (28,29).

Chloramphenicol acetyltransferase (CAT) has been used as an N-terminal tail for the purification of human atrial natriuretic factor (30,31). Fusion proteins containing various amounts of CAT were constructed, and the largest of these retained CAT activity, allowing purification by substrate-affinity chromatography. Although thrombin was successfully used to cleave the tail from the target protein at a genetically engineered site, two other methods of tail removal (enterokinase or BNPS–skatole cleavage) were more problematic.

TABLE 1

Tail Types

General type	Size	C- or N-terminus	Separation method	Ref.
1. Enzymes				
βGal	116 kDa	N, C	IS, IA	(8–27)
GST	26 kDa	N	IS (GSH)	(28, 29)
CAT	24 kDa	N	IS	(30, 31)
TrpE	27 kDa	N	HIC	(32)
2. Polypeptide-binding proteins				
SPA	14–31 kDa	N	I-IgG	(33–42)
SPG	28 kDa	C	I-albumin	(42, 74)
3. Carbohydrate-binding domains				
MBP	40 kDa	N	IS	(43–46)
SBD	119 aa	C	IS	(48, 49)
CBD_{CenA}	111 aa	N	IS	(50, 52)
CBD_{Cex}	128 aa	C	IS	(51–53)
4. Biotin-binding domain	8 kDa	N	I-avidin	(55)
5. Antigenic epitopes				
recA	144 aa	C	IA	(56)
Flag	8 aa	N	IA	(58, 59)
6. Charged amino acids				
Poly(Arg)	5–15 aa	C	IEC, pptn	(60–62, 65)
Poly(Asp)	5–16 aa	C	IEC, pptn, RME	(66–68)
Glutamate	1 aa	N	IEC	(63)
7. Poly(His) tails	1–9 aa	N, C	IMAC	(64, 71–74)
8. Other poly(amino acid) tails				
Poly(Phe)	11 aa	N	Phenyl-Superose	(75)
Poly(Cys)	4 aa	N	Thiopropyl-Sepharose	(75)

Note. Abbreviations used: aa, amino acid; βGal, β-galactosidase; CAT, chloramphenicol acetyltransferase; CBD, cellulose-binding domain; GSH, reduced glutathione; GST, glutathione-*S*-transferase; HIC, hydrophobic interaction chromatography; IA, immobilized antibody; IEC, ion-exchange chromatography; I-IgG, immobilized immunoglobulin G; IMAC, immobilized metal affinity chromatography; IS, immobilized substrate; MBP, maltose-binding protein; pptn, selective precipitation; RME, reverse micellar extraction; SBD, starch binding domain; SPA, staphylococcal protein A; SPG, streptococcal protein G.

Target proteins fused to the *E. coli TrpE* gene product can be stably expressed in large amounts but are often insoluble (7). Hummel *et al.* (32) have exploited the hydrophobicity of insoluble *TrpE* fusion proteins for their purification, using HPLC with a C_8 reversed-phase column and eluting with a 2-propanol gradient in formic acid.

2. *Polypeptide-Binding Proteins*

The immunoglobulin G (IgG)-binding domains of staphylococcal protein A (SPA) have a strong affinity ($K_D \approx 10^{-8}$ M) for the constant region F_c of mammalian IgGs and have been utilized extensively as N-terminal tails for purification on immobilized IgG columns (33–39). Vectors containing two IgG-binding E domains of SPA are available for either intracellular expression of fusion proteins in *E. coli* or, by including the SPA signal sequence, for periplasmic secretion from *E. coli* or extracellular excretion from *Staphylococcus aureus* (36,37,40). Secretion to the periplasm (or excretion to the culture medium) not only greatly simplifies purification, but also promotes proper folding and avoids proteolysis of the target protein. Insulin-like growth factor (IGF-I) was purified to the 95% level from *S. aureus* using a tail containing a pair of synthetic IgG-binding domains (36). The fusion protein, in clarified culture medium, was bound to an IgG-Sepharose column followed by elution with acetic acid. After lyophilization, the tail was cleaved at a unique hydroxylamine (Asn–Gly) cleavage site engineered to restore the authentic N-terminus of IGF-I. The tail was purified from the cleaved IGF-I by a second passage through the IgG column. A similar procedure was adopted for successful scale-up of IGF-I isolation from a 1000-liter fermentation (41). The SPA tail in some instances has been shown to protect fusion proteins from proteolysis (36). In some cases, removal of the tail by hydroxylamine cleavage has been inefficient (38), but a variety of different protease-specific cleavage sites are now available in various SPA vectors (37). A major application of SPA fusions which does not require tail removal is isolation of antibodies (35), a process that is facilitated by the SPA tail acting as an adjuvant.

Dual affinity fusion proteins with a pair of synthetic IgG-binding domains at the N-terminus and an albu-

min-binding domain (streptococcal protein G (SPG)) at the C-terminus have also been constructed for the isolation of extremely sensitive peptides in *E. coli* (42). Sequential affinity purification on immobilized IgG followed by purification on immobilized albumin greatly increased the recovery of full-length intervening target protein (human insulin-like growth factor II). Interestingly, the presence of the C-terminal tail stabilized the fusion protein, resulting in decreased proteolytic degradation. The presence of the larger, soluble tails at the N- and C-termini of the target also promoted intracellular solubility of the fusion protein.

3. *Carbohydrate-Binding Domains*

Several fusion protein recovery systems have been designed to take advantage of low-cost complex carbohydrates as adsorbents, utilizing mild conditions for both binding and elution. The periplasmic maltose-binding protein (MBP) of *E. coli* (370 residues) has been developed as an N-terminal tail for purification of target proteins either expressed intracellularly or secreted to the periplasm (43–45). Purification is based on the strong affinity of MBP to cross-linked amylose. After binding, fusion proteins can be eluted with maltose at a high yield (70%). An *E. coli* expression vector that includes a blood factor X_a site for removal of the tail after purification by repassage over the cross-linked amylose column is available (46). The tail could be cleaved under nondenaturing conditions from a fusion with paramyosin from *Dirofilaria imminitis,* but denaturation with urea was required to cleave the tail from a βGal fusion.

A wide variety of carbohydrases with independent domains that bind insoluble substrates (e.g., starch, cellulose) have been shown to have similar domain architecture (47). In these enzymes, the substrate-binding domain (located at either the N- or the C-terminus) is linked to a separate catalytic domain by a hinge of 50 or more hydroxyl-rich amino acids (and in some cases also proline). We have shown that a βGal fusion protein containing the starch-binding domain (SBD) of *Aspergillus awamori* glucoamylase at the C-terminus is functional when expressed in *E. coli* and could be purified from crude soluble cell extracts to a high level of purity by binding and elution from native starch granules under mild conditions (using boric acid–borax buffer, pH 8.2) (48). We have recently defined the optimum size of the domain for stability and activity in fusion proteins (119 aa) and shown that starch has a loading capacity for a βGal fusion protein with the 119 amino acid tail of 36.5 mg protein/g starch (49). The fusion protein can also be purified on a cross-linked amylose column, using maltose as an eluant (Z. Nikolov, personal communication).

Cellulose-binding domains (CBD) from two *Cellulomonas fimi* cellulases have been used to construct fusion proteins that can be purified by adsorption to and elution from cellulose (50,51). The CBD from *C. fimi* endoglucanase CenA is located at the N-terminus of the enzyme and has been used as an N-terminal fusion tail to purify alkaline phosphatase secreted to the *E. coli* periplasm (50). The CBD from *C. fimi* exoglucanase Cex is located at the C-terminus of the enzyme and has been used as a C-terminal fusion tail to purify β-glucosidase (52,53). Fusion proteins containing the CBD tails were bound to cellulose at low salt concentrations and could be eluted with either water or 8 M guanidine hydrochloride.

In a variation on carbohydrate binding, Taylor and Drickamer (54) have utilized the specific interaction between a lectin and a carbohydrate to design a purification fusion. They fused the carbohydrate-recognition domain (CRD) of the rat galactose-specific asialoglycoprotein receptor (rat hepatic lectin) to human placental alkaline phosphatase. The fusion protein was expressed in rat fibroblasts using a retroviral vector and in Sf9 insect cells using a baculovirus vector. The hybrid protein was secreted into the medium in both cases (due to the insulin signal sequence present in the constructs) and was recovered on a galactose-Sepharose column. The binding of CRD to galactose-Sepharose is Ca^{2+}-dependent, and the bound enzyme can be eluted under mild conditions by reducing the Ca^{2+} concentration in the buffer. Immobilized trypsin was used to cleave the alkaline phosphatase from CRD. Repurification in a galactose-Sepharose column removed the protease as well as the released CRD. This study, while basically similar to work in corresponding microbial systems, demonstrates the feasibility of the fusion tail approach in eukaryotic systems: the fusion protein can be secreted into the culture medium and recovered with high specificity by binding of the tail.

4. *Biotin-Binding Domain*

Cronan (55) has fused a peptide encoding a site for biotination both to the C-terminus of βGal and to the yeast HIS3 protein. The target proteins were biotinated *in vivo* upon expression in *E. coli* or yeast, allowing purification by affinity to immobilized avidin or streptavidin. The fusion protein tail contains 75 amino acids from the 1.3 S subunit of *Propionibacterium shermanii* transcarboxylase, which provides a functional site for *in vivo* biotination by intracellular biotin ligase. Biotin-binding fusion proteins can be isolated based either on their extremely high affinity to unmodified (tetrameric) avidin or streptavidin ($K_D = 10^{-15}$ M) or on their more moderate affinity to modified (monomeric) avidin ($K_D = 10^{-6}$–10^{-7} M). Recovery of proteins from unmodified avidin or streptavidin can be accomplished by cleaving the fusion with a site-specific protease or by boiling in SDS (a potential advantage for the recovery of membrane proteins). Proteins can be recovered in native

form from modified avidin columns by elution in mild buffers containing biotin. Very few bacterial or eukaryotic proteins are naturally biotinated, resulting in a high level of purification for fusion proteins containing the biotin-binding tail.

5. *Antigenic Epitopes*

An early development in immunoaffinity purification of fusion proteins was described by Krivi *et al.* (56). A soluble *recA*-somatostatin fusion protein was produced intracellularly in *E. coli* and purified with an immobilized anti-*recA* monoclonal antibody. Only a part of the *recA* polypeptide chain (including the antigenic determinant) was needed for efficient recovery of the hybrid protein.

Wigler and co-workers have also used epitope addition to purify a RAS-responsive adenylyl cyclase complex from yeast (57). Since adenylyl cyclase is difficult to purify, a short region coding for an antigenic peptide was fused to the 5′ end of the adenyl cyclase gene to produce an N-terminal fusion tail. The resulting hybrid protein, expressed in yeast, was recovered from the membrane fraction by immunoaffinity chromatography utilizing a Sepharose-bound monoclonal antibody raised against the tail. Competitive elution of the hybrid protein was achieved using a synthetic peptide comprising part of the original antigenic tail and resulted in a >700-fold purification of the protein and associated complex.

Hopp *et al.* (58) have designed an extremely hydrophilic (and therefore antigenic) eight amino acid tail (Asp–Tyr–Lys–Asp–Asp–Asp–Asp–Lys) that has been used successfully for one-step purification of recombinant target proteins by affinity to a calcium-dependent immobilized monoclonal antibody. The tail was designed for fusion at the N-terminus of secreted proteins, and expression vectors have been constructed for both *E. coli* and yeast, with the tail located between the secretion signal sequence and the authentic N-terminus of the target protein. The *E. coli* vector utilizes a *lac*-regulated *lpp* promoter and the *ompA* signal, while yeast vectors constructed utilize either the α-factor or ADH2 promoter and the α-factor signal sequence. The first four amino acids of the tail define an antigenic epitope that is recognized by a calcium-dependent monoclonal antibody (59). Elution from the column can therefore be achieved by lowering the effective calcium concentration of the column buffer. The last five amino acids of the tail specify an enterokinase cleavage site for removal of the tail from the target protein after purification if necessary. Although the presence of the tail did not affect the biological activity of the target proteins tested (interleukin 2, and granulocyte–macrophage colony-stimulating factor) cleavage of the tail was designed to allow production of authentic N-termini.

6. *Charged Amino Acids*

Positively charged poly(Arg) tails have been developed for recovery and purification of target proteins by ion-exchange chromatography (60-62). A five-residue poly(Arg) tail engineered at the C-terminus of recombinant human urogastrone was used to facilitate purification of the target protein to the 95% level. Following initial ion-exchange chromatography of the fusion protein, the tail was digested with carboxypeptidase B (specific for C-terminal Arg and Lys residues), and the target protein was repurified by ion-exchange chromatography. The two-step purification process resulted in a 44% yield of purified target protein. Tail removal by carboxypeptidase B can be adapted for target proteins that need to retain a C-terminal Arg or Lys by incorporating an adjacent protective amino acid that is resistant to digestion. The protective amino acid could then be removed by carboxypeptidase A, without digesting the authentic C-terminal Arg or Lys.

Dalbøge *et al.* (63) fused an N-terminal dipeptide tail containing a single charged residue (glutamate) to human growth hormone to promote purification by ion-exchange chromatography. The tail was designed to be removed by an exodipeptidase (DAP I) since the authentic N-terminal residue of the target is proline, which blocked further digestion. This strategy can be adapted to proteins without an N-terminal proline by addition of proline to the tail and use of a second peptidase (DAP IV). The glutamate residue in the tail of the fusion protein provided both a means of initial purification and a sensitive indicator by ion-exchange HPLC of incomplete tail removal after exopeptidase treatment.

Positively and negatively charged tails composed of multiple arginine or aspartic acid residues have also been used by Glatz and co-workers for recovery and purification by polyelectrolyte precipitation and other lower-cost charge-based methods. βGal fusion proteins containing 5–15 C-terminal poly(Arg) tails have been constructed for expression in *E. coli*. It was necessary to express the poly(Arg) fusion proteins in a strain lacking the outer membrane *ompT* protease in order to avoid excessive degradation of the positively charged tails, since this protease cleaves between two adjacent basic amino acids. Additional precautions had to be taken to avoid loss of expression of the poly(Arg) fusion proteins due to high rates of plasmid instability or reversion. This problem has also been noted by other workers for fusion proteins containing poly(Arg) tails (64). The poly(Arg) fusion proteins can be precipitated from crude *E. coli* extracts at a high level (compared to native βGal) with poly(acrylic acid) at a pH slightly above the isoelectric point. In contrast to poly(Asp) fusions (see below) to the same protein, it is not necessary to first remove the nucleic acids (65).

Unlike with positively charged tails, βGal fusion proteins containing poly(Asp) tails were neither proteolytically degraded nor unstably expressed (66,67). C-terminal tails of 5 and 11 aspartic acid residues promoted increasing precipitation of the target protein (βGal) by polyethyleneimine without affecting specific activity. A poly(Asp) acid tail of 16 residues, however, caused loss of specific activity and stability, without additional improvement in precipitation. Nucleic acids interfered with purification of the negatively charged fusion proteins from soluble cell extracts, but could be removed by pretreatment with nuclease followed by diafiltration. After nuclease treatment, the fusion protein with $poly(Asp)_{11}$ could be precipitated from soluble cell extracts with a fivefold increased selectivity compared to a βGal control (67). Negatively charged C-terminal tails have also been shown to promote polyethyleneimine precipitation of a secreted monomeric protein expressed by yeast (*Aspergillus* glucoamylase), resulting in very large separation factors (C. Ford, unpublished results). The negatively charged tails also enhanced the uptake of glucoamylase into a reversed-micellar system using the cationic surfactant, trioctylmethyl ammonium chloride (68).

7. *Poly(His) Tails*

Immobilized metal affinity chromatography (IMAC) is a powerful technique based on interactions of some proteins with immobilized transition metal ions. Although still not fully understood, the interactions depend on the formation of coordination complexes between transition metal ions (Ni^{2+}, Zn^{2+}, Cu^{2+}, or Co^{2+}) and the imidazole group of histidine residues or the sulfhydryl group of cysteine residues (69,70). The metal ions involved in the interactions can be immobilized for column chromatography on chelating solid supports.

Poly(His) tails ranging in length from two to six residues and fused to either the N- or the C-terminus of mouse dihydrofolate reductase (DHFR) were shown by Hochuli *et al.* (71) to promote purification by Ni^{2+}-IMAC. In chromatography using phosphate buffer, 90% of the DHFR with a C-terminal tail containing two histidines bound to the nickel chelate column and could be completely eluted with a decreasing pH gradient. DHFR containing six histidines bound to the column under these conditions, but only 10% could be eluted. When 6 M guanidine HCl was used in the loading and eluting buffers, only DHFR with six histidines at either the N- or the C-terminus bound to (and therefore could be eluted from) the column at a high (90%) level. In these studies, the C-terminal poly(His) tails were removed following recovery using carboxypeptidase A. Smith *et al.* (72) have also employed Ni^{2+}-IMAC to purify recombinant proinsulin using a dipeptide N-terminal (His–Trp) tail.

Ljungquist *et al.* (73) have used Zn^{2+}-IMAC to purify fusion proteins containing multiple copies of the peptide Ala–His–Gly–His–Arg–Pro. Fusion proteins containing four or eight copies of the peptide bound to the Zn^{2+}-IMAC column and could be eluted at pH 4.5–5.0. Fusion proteins containing two copies of the peptide (and presumably other background proteins) were unable to bind to the column. Plückthun and co-workers have also used Zn^{2+}-IMAC for one-step purification of an immunoglobulin G Fv heavy light chain hybrid with a C-terminal tail containing five histidine residues (64). The hybrid was secreted into the *E. coli* periplasm due to the presence of an N-terminal PhoA signal sequence. The C-terminal poly(His) residues apparently did not perturb the structure of the Fv fragment since hapten binding was unaffected by the tail. The tail variant $Pro(His_6)Pro$ bound to the column about as well as the His_5 tail; however, tails containing $(Gly–His)_3$ or $(His–Trp)_3His$ bound more weakly to the column. Poly(His) tails longer than five residues (His_9) or shorter (His_3) performed poorly in separations and the fusion proteins containing them were expressed at reduced levels.

Jansson *et al.* (74) have used a dual-affinity fusion system to recover target proteins (synthetic IgG-binding domains) that were extremely susceptible to proteolysis. In this case, the dual-affinity system utilized a C-terminal poly(His) tail together with an N-terminal albumin-binding domain (SPG). After release from the *E. coli* periplasm, the fusion proteins were purified by sequential chromatography on albumin-Sepharose and Zn^{2+}-chelating Sepharose and retained full target protein activity. As was found for other dual-affinity fusions (42), the presence of the C-terminal tail inceased the stability of the protein, resulting in decreased proteolytic degradation.

8. *Other Poly(Amino Acid) Tails*

Hydrophobic interactions have been used by Persson *et al.* (75) for the purification of βGal containing an N-terminal 11-residue poly(Phe) tail. The fusion protein was purified on a phenyl-Superose FPLC column, using an ethylene glycol gradient for elution of the fusion protein.

Persson *et al.* (75) have also purified *E. coli* galactokinase by engineering a four-residue poly(Cys) tail at the N-terminus and utilizing a thiopropyl-Sepharose column. The recombinant enzyme was eluted with 80% purity by addition of the reducing agent dithiothreitol.

CONSIDERATIONS FOR CHOOSING A FUSION TAIL SYSTEM

When choosing a fusion tail system for a specific application, several issues need to be considered:

1. Will the fusion protein be produced in a eukaryotic system? Potential advantages of a eukaryotic system

(especially for eukaryotic target proteins) include proper folding and post-translational modifications. It may be desirable to express therapeutic proteins in higher eukaryotic cells (rather than in yeast or fungi) in order to achieve more authentic glycosylation patterns.

2. If the fusion protein is produced in *E. coli*, will it be expressed intracellularly? Intracellular production results in much higher yields, but overexpressed heterologous proteins are often insoluble. If the target protein must be solubilized from inclusion bodies, it is important that the recovery system be able to operate in strong denaturing buffers. Other problems such as incorrect protein folding, lack of post-translational modification, and proteolysis may present themselves (see Cellular Location and Solubility of the Protein Product).

3. Will the target protein require stabilization for efficient heterologous expression? If so, it may be advantageous to use one of the larger *E. coli* enzymes such as βGal as a tail, or a dual-affinity tail system.

4. Will the fusion protein be secreted to the *E. coli* periplasm? Potential advantages of periplasmic (or extracellular) secretion are decreased proteolysis and decreased levels of background protein. Some tails are naturally secreted (e.g., protein A), while others are not (e.g., βGal) or may affect secretion levels (74).

5. Will the tail interfere with biological activity of the target? It has generally been thought that the smaller the tail the better, unless the fusion protein needs to be stabilized by the tail. Thus, an effort has been made to minimize the size of many tails in order to minimize possible interference with the target protein (and to increase the relative amount of target protein in total cellular output).

6. Will the tail be used as an assay tag to monitor the presence of the target protein? If so, the availability of tail-specific antibodies or the enzymatic properties of the tail will need to be considered.

7. Does the tail need to be removed after recovery? If so, the site for enzymatic or chemical cleavage of the tail must not also be found within the target protein (see Tail Cleavage).

8. Does the N-terminal Met of the target protein need to be processed? If so, this may be conveniently done concurrently with removal of an N-terminal tail.

9. Does a vector already exist that can be used for convenient construction of the fusion protein? If not, the specific elements for cloning such a vector will have to be assembled (see Genetic Engineering of Fusion Proteins).

10. Is scale-up desired? Not all fusion tail systems are equally practical for large-scale use (see Applications).

GENETIC ENGINEERING OF FUSION PROTEINS

In the past decade, the technical hurdles involved in genetically engineering fusion proteins have diminished greatly. This has resulted from the widespread availability of methodologies for DNA manipulation (76), the advent of commercial sources for molecular biology products, and improvements in every phase of gene technology: transformation of cells, site-directed mutagenesis, automated oligonucleotide synthesis, automated DNA sequencing, etc. While the genetic engineering of a complex fusion protein expression system is not trivial, it is feasible to construct in a modest amount of time almost any combination of components or features desired. Genes encoding many of the fusion tails discussed above are currently available in useful expression vectors that contain multiple restriction sites for insertion of the target gene and include regions encoding specific cleavage sites for removal of the tail. Specific changes in these vectors or addition of other components can be made in a straightforward manner by site-directed mutagenesis (77) or by inserting synthetic gene fragments or fragments obtained from other sources.

CELLULAR LOCATION AND SOLUBILITY OF THE PROTEIN PRODUCT

Two common problems in expressing foreign proteins are instability and inclusion body formation. Small eukaryotic proteins, like peptide hormones, are especially sensitive to host proteases due to incorrect folding or foreign topology. Sometimes this can be overcome by altering the location of the product from intracellular to secreted (or periplasmic), and thereby providing a less reducing environment for folding (78). This approach has been adopted in most expression systems (79–81). The envelope of bacteria like *E. coli*, however, is not devoid of proteolytic activities (78,82–85). In fact, sometimes secretion results in significant proteolysis (86). Other potential drawbacks of secretion are that the product often needs to be a naturally secreted protein; a secretion signal must be engineered at the 5′ end of the protein coding sequence to direct the target protein across the cell membrane; and expression levels are usually not very high. Alternative approaches to stabilizing intracellularly unstable recombinant proteins include using dual tails (42,74), or fusion to certain proteins, like ubiquitin, which stabilizes the target protein (87).

Many foreign proteins expressed intracellularly at high levels in *E. coli* and yeasts are produced as insoluble aggregates, often called refractile or inclusion bodies (1,88). Inclusion body formation is not limited to intracellular expression but can take place in the *E. coli* periplasm as well (89). In inclusion bodies, proteins are protected from proteolytic attack (7,90), but they exist in a partially unfolded state and usually cannot be solubilized without the use of strong chaotropic agents such as guanidine hydrochloride or urea. After this treatment the proteins are completely denatured and must be re-

folded during or after removal of the denaturing agent. Fusions to *TrpE* and to a lesser extent βGal and GST may result in inclusion body formation.

TAIL CLEAVAGE

The question of whether to remove the purification tail or not is essentially dependent on the end-use of the target protein. For lab-scale characterization of a protein that has previously been difficult to obtain in sufficient quantities, it may be possible to leave the tail on, after initial demonstration that the tail does not interfere with the biological function of the target protein. If interference is encountered, removal of the tail may be necessary in order to obtain reliable results. For antibody production, it may be desirable to leave the fusion tail on as a carrier to promote antigenicity. In some applications, a slight or even considerable loss of activity can be justified by the ease of purification and assayability provided by the fusion moiety. Target proteins in this category may include industrial enzymes, diagnostic proteins, and enzymes that are to be immobilized. In the latter case the tail would be designed to promote simultaneous purification and immobilization by binding to a solid support. For pharmaceutical applications, precise removal of the fusion tail is usually desired in order to achieve absolute product authenticity. This requirement poses a serious bottleneck in the usage of fusion proteins on a production scale, since as yet there are no ideal methods for tail removal.

Two general classes of methodology are available for cleavage of the fusion proteins after initial purification: enzymatic methods and chemical methods. The basic methodology is derived from protein structural studies, especially peptide mapping and sequencing (see (91–93)). Updated lists of cleavage methods have been summarized previously by several authors (1–4,6). The essential features of the various cleavage methods are specificity, efficiency, and working (incubation) conditions. The problem is that these requirements tend to be mutually exclusive. Chemical methods generally have high efficiency but are usually rather nonspecific and definitely too harsh in most cases where biological activity is to be retained. Chemical cleavage sites typically are specified by only one or two residues (e.g., methionine as a site for cyanogen bromide cleavage) and are, therefore, often found within the target protein.

Enzymatic methods include exopeptidases with differing specificities and endopeptidases, some of which have recognition sites of up to five residues. Enzymatic methods can be employed under mild conditions, making their use the method of choice in most cases. Examples of very specific endopeptidases that have been used successfully include enterokinase (58), collagenase (26), the clotting enzyme factor X_a (6,46) and thrombin (32). Unfortunately, all of these suffer from fairly inefficient cleavage and, most of all, the presence of contaminating proteolytic activities (especially in commercial preparations). Since peptide cleavages for the removal of tails often occur at the C-terminal end of the cleavage site, authentic target proteins in most cases will be recoverable only if the tail (and cleavage site) is fused to the N-terminus of the target protein.

APPLICATIONS

Fusion tail technology can be applied to recovery and purification problems at both a lab scale and an industrial scale. The laboratory scale provides the most immediate applications for the majority of fusion recovery systems. At this scale the wide variety of specific tail interactions available for use is a major advantage, while regulatory concerns over ligand leakage and the cost concerns of affinity matrix production and column life are less important.

Lab-Scale Applications

Fusion tail technology at the lab scale finds application (1) in the routine recovery of proteins expressed by cDNA, especially for antibody production; (2) in the recovery, assay, and characterization of recombinant proteins being developed for therapeutic or industrial use; (3) for the recovery of "difficult" proteins from any source that have previously evaded in depth characterization due to low expression or low purification yields; and (4) for the isolation and purification of a third protein that has affinity to the fusion protein.

As an example of the latter category, Flanagan and Leder (94) have described a general approach for the isolation of ligands to receptors with unknown function. They fused a human placental alkaline phosphatase tail to the extracellular domain of the mouse c-*kit* proto-oncogene cell surface receptor, replacing the intracellular and hydrophobic transmembrane domains of the receptor. The fusion protein was expressed and secreted from a cultured mouse cell line, and could be recovered and purified by anti-alkaline phosphatase immunoaffinity chromatography for use in identifying ligands that bind to the receptor. Similarly, fusion proteins have been constructed to facilitate detection (95) and purification (96) of specific antibodies from polyclonal antibody mixtures.

Fusion proteins can also be used to isolate and purify DNA-binding proteins. Levens and Howley (97) describe a general approach for the isolation of sequence-specific DNA-binding proteins from complex mixtures. In this scheme, the DNA sequence of interest is first cloned adjacent to the *lac* operator DNA segment. This hybrid DNA is then bound to a fusion protein consisting of the *lac* repressor and βGal, and the complex is bound, via the βGal moiety, to immobilized anti-βGal. The bound complex is incubated with cell extracts in order

to isolate DNA-binding proteins specific for the DNA fragment cloned adjacent to the *lac* operator. The DNA-binding protein complex can be released with isopropyl-β-D-thiogalactopyranoside for characterization and analysis of the isolated protein (or proteins if multiple protein-binding sites exist within a short stretch of DNA).

Another example of the use of fusion proteins is the purification by receptor-affinity chromatography of an interleukin 2-*Pseudomonas* exotoxin fusion protein, a potentially cytotoxic agent for interleukin-2 receptor-bearing cells (98). The term receptor-affinity chromatography was coined because immobilized interleukin-2 receptors were used as the affinity matrix.

Fusion proteins often serve more than one purpose. As an example, Krozowski (9) has used a glutathione-*S*-transferase tail both to promote purification of a rat whey acidic protein (WAP) and to provide sites (tyrosine residues) for radioiodination in order to develop an efficient WAP radioimmunoassay.

Larger-Scale Applications

At larger scales, where the cost and scale-up limitations of chromatographic separations are more troubling, it is worth pursuing alternatives to affinity adsorption, especially for larger volume proteins. Nonadsorptive methods (e.g., precipitation) have been applied with high selectivity when the target protein has its own distinctive site for interaction (affinity interactions being the most obvious candidate) with the separating agent. Hence, using fusion tail technology, one could take advantage of the selectivities demonstrated for affinity precipitation based on any of a variety of peptide–ligand interactions. Examples are biotin–avidin binding with the biotin as a ligand on hydrophobic precipitants (100); NAD binding to NADH–dehydrogenase (101) where the oligomeric protein was precipitated with a multi-NAD precipitant; and binding to ligand-derivatized polymers whose solubility can be manipulated. The last case is illustrated by the recovery of trypsin using a polymer-bound inhibitor (102). A drawback to affinity precipitation is that the yields obtained can be very sensitive to the stoichiometric ratios of protein and precipitant.

Such a strategy is not limited to precipitation. Similar affinity interactions have been used in several other ways: to enhance partitioning in two-phase aqueous extraction by binding ligands to polyethylene oxide (103); to enhance reversed-micellar extraction by coupling ligands to the cosurfactant (104); and to control rejection behavior in ultrafiltration (105). While these operations are convenient to use at larger scales, there are still the higher costs and regulatory concerns associated with the affinity reagents. Little work has yet been reported on the pairing of possible fusion tails with these nonadsorptive methods.

A number of the drawbacks associated with tails requiring affinity ligands can be avoided by tails designed to enhance properties such as protein charge or hydrophobicity, which enable purifications based on traditional reagents. The previously cited use of poly(Asp) and poly(Arg) tails enabled the recovery of fusion proteins using polyelectrolyte precipitation (66,67), ion-exchange (61), aqueous-phase partitioning (106), and reversed-micellar extraction (68). The same electrostatic interactions appear to be sufficient to control rejection behavior in ultrafiltration (107). An added benefit of the pairing of fusion proteins with these separation methods is that a better understanding of the mechanisms of separation may be obtained.

In addition to choosing a method/tail pair, one must also choose where in the multistep downstream processing sequence to use it. The ideal of a highly selective one-step separation performed on a crude extract is not likely to be realized very often in practice. This is the promise of affinity adsorption, but considerations of column fouling and ligand leakage will usually require some additional treatment. Hoare (108) considered the advisibility of using costly affinity resins only where full advantage could be taken of their potential selectivity. In a study of separation schemes from the literature, advantages to some preliminary separation were found.

There are two barriers to the use of purification fusions for therapeutics that have nothing to do with shortcomings in purification potential. The first is the need to introduce their use early in the product development process. In that way, the use can be part of the product approval process from the beginning. The second barrier is the need to remove the tail so that the final product is the unaltered target protein. Clearly it is advantageous to have an application where the fusion is acceptable and the tail does not diminish activity.

CONCLUSION

At present, it seems reasonable to develop purification fusion approaches in the following cases: (1) for characterization of proteins at the lab scale, when authentic proteins prove to be too difficult (unstable or otherwise) to purify; (2) for production of antibodies against unstable or otherwise problematic proteins; (3) for development of (immuno)assays or other analytic tools for the target protein; (4) for use in dual-purpose applications such as covalent modification or immobilization; (5) for production of extremely valuable pharmaceutical products, when other current approaches (maximization of expression, minimization of cell culture costs, etc.) are not adequate; and (6) for production of lower-cost industrial enzymes, especially if the tail

can be left on and there are real savings in separation costs by a reduction of the number of steps.

ACKNOWLEDGMENTS

We are grateful to Deborah Stachon for help in preparing the manuscript. This is Journal Paper J-14503 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Projects 2820 and 2822.

REFERENCES

1. Marston, F. A. O. (1986) The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem. J.* **240**, 1–12.
2. Sassenfeld, H. M. (1990) Engineering proteins for purification. *Trends Biotechnol.* **8**, 88–93.
3. Uhlén, M., and Moks, T. (1990) Gene fusions for purpose of expression: An introduction, *in* "Methods in Enzymology" (Goeddel, D. V., Ed.), Vol. 185, pp. 129–143, Academic Press, San Diego.
4. Hammond, P. M., Atkinson, T., Sherwood, R. F., and Scawen, M. D. (1991) Manufacturing new-generation proteins. Part I. The technology. *BioPharm* **4**, 16–23.
5. Sherwood, R. (1991) Protein fusions: Bioseparation and application. *Trends Biotechnol.* **9**, 1–3.
6. Nagai, K., and Thøgersen, H. C. (1987) Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*, *in* "Methods in Enzymology" (Wu, R., and Grossman, L., Eds.), Vol. 153, pp. 461–481, Academic Press, San Diego.
7. Yansura, D. G. (1990) Expression as *trpE* fusion, *in* "Methods in Enzymology" (Goeddel, D. V., Ed.), Vol. 185, pp. 161–167, Academic Press, San Diego.
8. Germino, J., and Bastia, D. (1984) Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. *Proc. Natl. Acad. Sci. USA* **81**, 4692–4696.
9. Ullmann, A. (1984) One-step purification of hybrid proteins which have β-galactosidase activity. *Gene* **29**, 27–31.
10. Offensperger, W., Wahl, S., Neurath, A. R., Price, P., Strick, N., Kent, S. B. H., Christman, J. K., and Acs, G. (1985) Expression in *Escherichia coli* of a cloned DNA sequence encoding the pre-S2 region of hepatitis B virus. *Proc. Natl. Acad. Sci. USA* **82**, 7540–7544.
11. Struck, D. K., Maratea, D., and Young, R. (1985) Purification of hybrid β-galactosidase proteins encoded by φX174 EφlacZ and *Escherichia coli* prlAφlacZ: A general method for the isolation of lacZ fusion polypeptides produced in low amounts. *J. Mol. Appl. Genet.* **3**, 18–25.
12. Hirel, P.-H., Lévêque, F., Mellot, P., Dardel, F., Panvert, M., Mechulam, Y., and Fayat, G. (1988) Genetic engineering of methionyl-tRNA synthetase: *In vitro* regeneration of an active synthetase by proteolytic cleavage of a methionyl-tRNA synthetase-β-galactosidase chimeric protein. *Biochimie* **70**, 773–782.
13. Nielsen, O. J., Costa-Giomi, P., Weinmann, R., Erslev, A. J., and Caro, J. (1988) Erythropoietin-β-D-galactosidase. The generation, purification and use of a fusion protein. *J. Immunol. Methods* **111**, 1–9.
14. Rüther, U., and Müller-Hill, B. (1983) Easy identification of cDNA clones. *EMBO J.* **2**, 1791–1794.
15. Bartels, D., Thompson, R. D., and Rothstein, S. (1985) Synthesis of a wheat storage protein subunit in *Escherichia coli* using novel expression vectors. *Gene* **35**, 159–167.
16. Scholtissek, S., and Grosse, F. (1988) A plasmid vector system for the expression of a triprotein consisting of β-galactosidase, a collagenase recognition site and a foreign gene product. *Gene* **62**, 55–61.
17. Markmeyer, P., Rühlmann, A., Englisch, U., and Cramer, F. (1990) The pAX plasmids: New gene-fusion vectors for sequencing, mutagenesis and expression of proteins in *Escherichia coli*. *Gene* **93**, 129–134.
18. Scherf, A., Mattei, D., and Schreiber, M. (1990) Parasite antigens expressed in *Escherichia coli*; a refined approach for epidemiological analysis. *J. Immunol. Methods* **128**, 81–87.
19. Itakura, K., Hirose, T., Crea, R., and Riggs, A. D. (1977) Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* **198**, 1056–1063.
20. Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., and Riggs, A. D. (1979) Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc. Natl. Acad. Sci. USA* **76**, 106–110.
21. Shine, J., Fettes, I., Lan, N. C. Y., Roberts, J. L., and Baxter, J. D. (1980) Expression of cloned β-endorphin gene sequences by *Escherichia coli*. *Nature* **285**, 456–461.
22. Stanley, K. K., and Luzio, J. P. 1984) Construction of a new family of high efficiency bacterial expression vectors: Identification of cDNA clones coding for human liver proteins. *EMBO J.* **3**, 1429–1434.
23. Steers, E., Jr., Cuatrecasas, P., and Pollard, H. B. (1971) The purification of β-galactosidase from *Escherichia coli* by affinity chromatography. *J. Biol. Chem.* **246**, 196–200.
24. Shuman, H. A., Silhavy, T. J., and Beckwith, J. R. (1980) Labeling of proteins with β-galactosidase by gene fusion: identification of a cytoplasmic membrane component of the *Escherichia coli* maltose transport system. *J. Biol. Chem.* **255**, 168–174.
25. Harris, T. J. (1983) *in* "Genetic Engineering" (Williamson, R., Ed.), Vol. 4, pp. 127–185, Academic Press, London.
26. Hanada, K., Yamato, I., and Anraku, Y. (1988) Purification and reconstitution of *Escherichia coli* proline carrier using a site specifically cleavable fusion protein. *J. Biol. Chem.* **263**, 7181–7185.
27. Hickman, R. K., McMurry, L. M., and Levy, S. B. (1990) Overproduction and purification of the Tn*10*-specified inner membrane tetracycline resistance protein Tet using fusions to β-galactosidase. *Mol. Microbiol.* **4**, 1241–1251.
28. Smith, D. B., and Johnson, K. S. (1988) One-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31–40.
29. Guan, K., and Dixon, J. E. (1991) Eukaryotic proteins expressed in *Escherichia coli*: An improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. *Anal. Biochem.* **192**, 262–267.
30. Dykes, C. W., Bookless, A. B., Coomber, B. A., Noble, S. A., Humber, D. C., and Hohden, A. N. (1988) Expression of atrial natriuretic factor as a cleavable fusion protein with chloramphenicol acetyltransferase in *Escherichia coli*. *Eur. J. Biochem.* **174**, 411–416.
31. Knott, J. A., Sullivan, C. A., and Weston, A. (1988) The isolation and characterization of human atrial natriuretic factor produced as a fusion protein in *Escherichia coli*. *Eur. J. Biochem.* **174**, 405–410.
32. Hummel, M., Herbst, H., and Stein, H. (1989) Gene synthesis, expression in *Escherichia coli* and purification of immunoreactive human insulin-like growth factors I and II: Application of a modified HPLC separation technique for hydrophobic proteins. *Eur. J. Biochem.* **180**, 555–561.
33. Uhlén, M., Nilsson, B., Guss, B., Lindberg, M., Gatenbeck, S., and Philipson, L. (1983) Gene fusion vectors based on the gene for staphylococcal protein A. *Gene* **23**, 369–378.

34. Nilsson, B., Abrahmsén, L., and Uhlén, M. (1985) Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors. *EMBO J.* **4**, 1075–1080.

35. Löwenadler, B., Nilsson, B., Abrahmsén, L., Moks, T., Ljungqvist, L., Holmgren, E., Paleus, S., Josephson, S., Philipson, L., and Uhlén, M. (1986) Production of specific antibodies against protein A fusion proteins. *EMBO J.* **5**, 2393–2398.

36. Moks, T., Abrahmsén, L., Holmgren, E., Bilich, M., Olsson, A., Uhlén, M., Pohl, G., Sterky, C., Hultberg, H., Josephson, S., Holmgren, A., Jörnvall, H., and Nilsson, B. (1987) Expression of human insulin-like growth factor I in bacteria: Use of optimized gene fusion vectors to facilitate protein purification. *Biochemistry* **26**, 5239–5244.

37. Nilsson, B., and Abrahmsén, L. (1990) Fusions to staphylococcal protein A, *in* "Methods in Enzymology" (Goeddel, D. V., Ed.), Vol. 185, pp. 144–161, Academic Press, San Diego.

38. Boutelje, J., Karlström, A. R., Hartmanis, M. G. N., Holmgren, E., Sjögren, A., and Levine, R. L. (1990) Human immunodeficiency viral protease is catalytically active as a fusion protein: Characterization of the fusion and native enzymes produced in *Escherichia coli*. *Arch. Biochem. Biophys.* **283**, 141–149.

39. Stringer, K. F., Ingles, C. J., and Greenblatt, J. (1990) Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* **345**, 783–786.

40. Abrahmsén, L., Moks, T., Nilsson, B., Hellman, U., and Uhlén, M. (1985) Analysis of signals for secretion in the staphylococcal protein A gene. *EMBO J.* **4**, 3901–3906.

41. Moks, T., Abrahmsén, L., Osterlöf, B., Josephson, S., Östling, M., Enfors, S.-O., Persson, I., Nilsson, B., and Uhlén, M. (1987) Large-scale affinity purification of human insulin-like growth factor I from culture medium of *Escherichia coli*. *Bio/Technol.* **5**, 379–382.

42. Hammarberg, B., Nygren, P.-A., Holmgren, E., Elmblad, A., Tally, M., Hellman, U., Moks, T., and Uhlén, M. (1989) Dual affinity fusion approach and its use to express recombinant human insulin-like growth factor II. *Proc. Natl. Acad. Sci. USA* **86**, 4367–4371.

43. Bedouelle, H., and Duplay, P. (1988) Production in *Escherichia coli* and one-step purification of bifuctional hybrid proteins which bind maltose. Export of the Klenow polymerase into the periplasmic space. *Eur. J. Biochem.* **171**, 541–549.

44. Guan, C., Li, P., Riggs, P. D., and Inouye, H. (1988) Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* **67**, 21–30.

45. Blondel, A., and Bedouelle, H. (1990) Export and purification of a cytoplasmic dimeric protein by fusion to the maltose-binding protein of *Escherichia coli*. *Eur. J. Biochem.* **193**, 325–330.

46. Maina, C. V., Riggs, P. D., Grandea, A. G., III, Slatko, B. E., Moran, L. S., Tagliamonte, J. A., McReynolds, L. A., and di Guan, C. (1988) An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene* **74**, 365–373.

47. Knowles, J., Lehtovaara, P., and Teeri, T. (1987) Cellulase families and their genes. *Trends. Biotechnol.* **5**, 255–261.

48. Chen, L., Ford, C., and Nikolov, Z. (1991) Adsorption to starch of a β-galactosidase fusion protein containing the starch-binding region of *Aspergillus* glucoamylase. *Gene* **99**, 121–126.

49. Chen, L., Ford, C., Kusnadi, A., and Nikolov, Z. (1991) Improved adsorption to starch of a β-galactosidase fusion protein containing the starch-binding domain from *Aspergillus* glucoamylase. *Biotechnol. Prog.* **7**, 225–229.

50. Greenwood, J. M., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1989) Fusion to an endoglucanase allows alkaline phosphatase to bind to cellulose. *FEBS Lett.* **244**, 127–131.

51. Ong, E., Greenwood, J. M., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1989) The cellulose-binding domains of cellulases: Tools for biotechnology. *Trends Biotechnol.* **7**, 239–243.

52. Ong, E., Gilkes, N. R., Warren, R. A. J., Miller, R. C., Jr., and Kilburn, D. G. (1989) Enzyme immobilization using the cellulose-binding domain of a *Cellulomonas fimi* exoglucanase. *Bio/Technol.* **7**, 604–607.

53. Ong, E., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. (1991) Enzyme immobilization using a cellulose-binding domain: Properties of a β-glucosidase fusion protein. *Enzyme Microb. Technol.* **13**, 59–65.

54. Taylor, M. E., and Drickamer, K. (1991) Carbohydrate-recognition domains as tools for rapid purification of recombinant eukaryotic proteins. *Biochem. J.* **274**, 575–580.

55. Cronan, J. E., Jr. (1990) Biotination of proteins *in vivo*; a post-translational modification to label, purify, and study proteins. *J. Biol. Chem.* **265**, 10,327–10,333.

56. Krivi, G. G., Bittner, M. L., Rowold, E., Jr., Wong, E. Y., Glenn, K. C., Rose, K. S., and Tiemeier, D. C. (1985) Purification of *recA*-based fusion proteins by immunoadsorbent chromatography: Characterization of a major antigenic determinant of *Escherichia coli recA* protein. *J. Biol. Chem.* **260**, 10,263–10,267.

57. Field, J., Nikawa, J.-I., Broek, D., MacDonald, B., Rodgers, L., Wilson, I. A., Lerner, R. A., and Wigler, M. (1988) Purification of a *RAS*-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. *Mol. Cell. Biol.* **8**, 2159–2165.

58. Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerretti, D. P., Urdal, D. L., and Conlon, P. J. (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technol.* **6**, 1204–1210.

59. Prickett, K. S., Amberg, D. C., and Hopp, T. P. (1989) A calcium-dependent antibody for identification and purification of recombinant proteins. *BioTechniques* **7**, 580–589.

60. Smith, J. C., Derbyshire, R. B., Cook, E., Dunthorne, L., Viney, J., Brewer, S. J., Sassenfeld, H. M., and Bell, L. D. (1984) Chemical synthesis and cloning of a poly(arginine)-coding gene fragment designed to aid polypeptide purification. *Gene* **32**, 321–327.

61. Sassenfeld, H. M., and Brewer, S. J. (1984) A polypeptide fusion designed for the purification of recombinant proteins. *Bio/Technol.* **2**, 76–81.

62. Brewer, S. J., and Sassenfeld, H. M. (1985) The purification of recombinant proteins using C-terminal polyarginine fusions. *Trends Biotechnol.* **3**, 119–122.

63. Dalbøge, H., Dahl, H.-H. M., Pedersen, J., Hansen, J., and Christensen, T. (1987) A novel enzymatic method for production of authentic hGH from an *Escherichia coli* produced hGH-precursor. *Bio/Technol.* **5**, 161–164.

64. Skerra, A., Pfitzinger, I., and Plückthun, A. (1991) The functional expression of antibody Fv fragments in *Escherichia coli*: Improved vectors and a generally applicable purification technique. *Bio/Technol.* **9**, 273–278.

65. Suominen, I., Forney, C., Luther, J., Niederauer, M., and Glatz, C. (1990) "Enhanced Selectivity in Bulk Separations Using Charged-Peptide Fusions," paper presented at the Engineering Foundation Conference on Recovery of Bioproducts, St. Petersburg Beach, FL, May 1990.

66. Zhao, J., Ford, C. F., Glatz, C. E., Rougvie, M. A., and Gendel, S. M. (1990) Polyelectrolyte precipitation of β-galactosidase fu-

sions containing poly-aspartic acid tails. *J. Biotechnol.* **14**, 273–284.

67. Parker, D. E., Glatz, C. E., Ford, C. F., Gendel, S. M., Suominen, I., and Rougvie, M. A. (1990) Recovery of a charged-fusion protein from cell extracts by polyelectrolyte precipitation. *Biotechnol. Bioeng.* **36**, 167–175.
68. Forney, C. E., and Glatz, C. E. (1991) "Behavior of Charged Fusion Proteins in Reversed Micellar Extraction," paper presented at the 65th ACS Colloid and Surface Science Symposium, Norman, OK, June 17–19, 1991.
69. Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598–599.
70. Hutchens, T. W., and Yip, T.-T. (1990) Protein interactions with immobilized transition metal ions: Quantitative evaluations of variations in affinity and binding capacity. *Anal. Biochem.* **191**, 160–168.
71. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., and Stüber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technol.* **6**, 1321–1325.
72. Smith, M. C., Furman, T. C., Ingolia, T. D., and Pidgeon, C. (1988) Chelating peptide-immobilized metal ion affinity chromatography. *J. Biol. Chem.* **263**, 7211–7215.
73. Ljungquist, C., Breitholtz, A., Brink-Nilsson, H., Moks, T., Uhlén, M., and Nilsson, B. (1989) Immobilization and affinity purification of recombinant proteins using histidine peptide fusions. *Eur. J. Biochem.* **186**, 563–569.
74. Jansson, B., Palmcrantz, C., Uhlén, M., and Nilsson, B. (1989) A dual-affinity gene fusion system to express small recombinant proteins in a soluble form: Expression and characterization of protein A deletion mutants. *Protein Eng.* **2**, 555–561.
75. Persson, M., Bergstrand, M. G., Bülow, L., and Mosbach, K. (1988) Enzyme purification by genetically attached polycysteine and polyphenylalanine affinity tails. *Anal. Biochem.* **172**, 330–337.
76. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Manual," 2nd ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
77. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82**, 488–492.
78. Hammarberg, B., Moks, T., Tally, M., Elmblad, A., Holmgren, E., Murby, M., Nilsson, B., Josephson, S., and Uhlén, M. (1990) Differential stability of recombinant human insulin-like growth factor II in *Escherichia coli* and *Staphylococcus aureus*. *J. Biotechnol.* **14**, 423–438.
79. Koller, K.-P., Reiß, G., Sauber, K., Uhlmann, E., and Wallmeier, H. (1989) Recombinant *Streptomyces lividans* secretes a fusion protein of tendamistat and paroinsulin. *Bio/Technol.* **7**, 1055–1059.
80. Stetler, G. L., Forsyth, C., Gleason, T., Wilson, J., and Thompson, R. C. (1989) Secretion of active, full- and half-length human secretory leukocyte protease inhibitor by *Saccharomyces cerevisiae*. *Bio/Technol.* **7**, 55–60.
81. Wang, L.-F., Wong, S.-L., Lee, S.-G., Kalyan, N. K., Hung, P. P., Hilliker, S., and Doi, R. H. (1988) Expression and secretion of human atrial natriuretic α-factor in *Bacillus subtilis* using the subtilisin signal peptide. *Gene* **69**, 39–47.
82. Baneyx, F., and Georgiou, G. (1990) In vivo degradation of secreted fusion proteins by the *Escherichia coli* outer membrane protease OmpT. *J. Bacteriol.* **172**, 491–494.
83. Gottesman, S. (1989) Genetics of proteolysis in *Escherichia coli*. *Annu. Rev. Genet.* **23**, 163–198.
84. Lazdunski, A. M. (1989) Peptidases and proteases of *Escherichia coli* and *Salmonella typhimurium*. *FEMS Microbiol. Rev.* **63**, 265–276.
85. Strauch, K. L., Johnson, K., and Beckwith, J. (1989) Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature. *J. Bacteriol.* **171**, 2689–2696.
86. Gentz, R., Kuys, Y., Zwieb, C., Taatjes, D., Taatjes, H., Bannwarth, W., Stueber, D., and Ibrahimi, I. (1988) Association of degradation and secretion of three chimeric polypeptides in *Escherichia coli*. *J. Bacteriol.* **170**, 2212–2220.
87. Butt, T. R., Jonnalagadda, S., Monia, B. P., Sternberg, E. J., Marsh, J. A., Stadel, J. M., Ecker, D. J., and Crooke, S. T. (1989) Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86**, 2540–2544.
88. Kane, J. F., and Hartley, D. L. (1988) Formation of recombinant protein inclusion bodies in *Escherichia coli*. *Trends Biotechnol.* **6**, 95–101.
89. Bowden, G. A., and Georgiou, G. (1990) Folding and aggregation of β-lactamase in the periplasmic space of *Escherichia coli*. *J. Biol. Chem.* **265**, 16,760–16,766.
90. Rhee, H. J., Lee, Y. I., and Yang, K. H. (1990) High-level expression of human insulin-like growth factor II in *Escherichia coli*. *J. Biotechnol.* **13**, 293–304.
91. Fontana, A., and Gross, E. (1986) Fragmentation of polypeptides by chemical methods, *in* "Practical Protein Chemistry—A Handbook" (Darbre, A., Ed.), pp. 67–120, Wiley, New York.
92. Ward, C. W. (1986) Carboxyl terminal sequence analysis, *in* "Practical Protein Chemistry—A Handbook" (Darbre, A., Ed.), pp. 191–525, Wiley, New York.
93. Wilkinson, J. M. (1986) Fragmentation of polypeptides by enzymic methods, *in* "Practical Protein Chemistry—A Handbook" (Darbre, A., Ed.), pp. 121–148, Wiley, New York.
94. Flanagan, J. G., and Leder, P. (1990) The *kit* ligand: A cell surface molecule altered in steel mutant fibroblasts. *Cell* **63**, 185–194.
95. Kwang, H. S., Barr, P. J., Sabin, E. A., Sujipto, S., Marx, P. A., Power, M. D., Bathurst, I. C., and Pedersen, N. C. (1988) Simian retrovirus-D serotype 1 (SRV-1) envelope glycoproteins gp70 and gp20: Expression in yeast cells and identification of specific antibodies in sera from monkeys that recovered from SRV-1 infection. *J. Virol.* **62**, 1774–1780.
96. Taatjes, D. J., Roth, J., Weinstein, J., and Paulson, J. C. (1988) Post-golgi apparatus localization and regional expression of rat intestinal sialytransferase detected by immunoelectron microscopy with polypeptide epitope-purified antibody. *J. Biol. Chem.* **263**, 6302–6309.
97. Levens, D., and Howley, P. M. (1985) Novel method for identifying sequence-specific DNA-binding proteins. *Mol. Cell. Biol.* **5**, 2307–2315.
98. Bailon, P., Weber, D. V., Gately, M., Smart, J. E., Lorberboum-Galski, H., Fitzgerald, D., and Pastan, I. (1988) Purification and partial characterization of an interleukin-2-*Pseudomonas* exotoxin fusion protein. *Bio/Technol.* **6**, 1326–1329.
99. Krozowski, Z. (1989) Generation of polyclonal antibodies against purified rat whey acidic proteins and the synthesis of a tracer fusion protein suitable for use in radioimmunoassays. *Mol. Cell. Endocrinol.* **63**, 15–22.
100. Guzman, R. Z., Kilpatrick, P. K., and Carbonell, R. G. (1990) Affinity precipitation of avidin by using ligand-modified surfactants, *in* "Downstream Processing and Bioseparations" (Hamel, J.-F., Hunter, J. B., and Sikdar, S. K., Eds.), ACS Symposium Series, Vol. 419, pp. 212–236, Amer. Chem. Soc., Washington, DC.

101. Larsson, P. O., and Mosbach, K. (1979) Affinity precipitation of enzymes. *FEBS Lett.* **98**, 333–338.

102. Schneider, M., Guillot, C., and Lamy, B. (1981) The affinity precipitation technique: Application to the isolation and purification of trypsin from bovine pancreas. *Ann. N.Y. Acad. Sci.* **369**, 257–263.

103. Johansson, G., Kopperschlager, G., and Albertsson, P. A. (1983) Affinity partitioning of phosphofructokinase from baker's yeast using polymer-bound Cibacron Blue F3G-A. *Eur. J. Biochem.* **131**, 589–594.

104. Woll, J. M., Hatton, T. A., and Yarmush, M. L. (1989) Bioaffinity separations using reversed micellar extraction. *Biotechnol. Prog.* **5**, 57–62.

105. Mattiasson, B., and Ling, T. G. I. (1986) Ultrafiltration affinity purification: A process for large-scale biospecific separations, *in* "Membrane Separations in Biotechnology" (McGregor, W. C., Ed.), pp. 99–114, Dekker, New York.

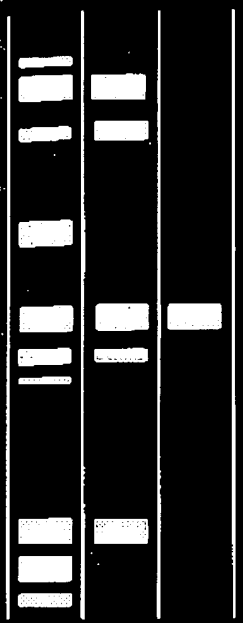
106. Luther, J. R., and Glatz, C. E. (1990) "Enhanced Partitioning in Aqueous Two-Phase Systems Using Genetically Engineered β-Galactosidase," poster presented at the AIChE Annual Meeting, Chicago, IL, November 1990.

107. Bozzano, A. G., and Glatz, C. E. (1991) Separation of proteins from polyelectrolytes by ultrafiltration. *J. Membr. Sci.* **55**, 181–198.

108. Hoare, M. (1990) "The Assessment of Chromatography Procedures for Protein Purification" paper presented at the Engineering Foundation Conference on Recovery of Bioproducts, St. Petersburg Beach, FL, May 1990.

Volume 2, Numbers 2/3
April/June 1991

PEXPEJ
ISSN 1046-5928



Protein Expression and Purification

ACADEMIC PRESS, INC. *Harcourt Brace Jovanovich, Publishers*
San Diego New York Boston London Sydney Tokyo Toronto

Protein Expression and Purification

Volume 2, Numbers 2/3, April/June 1991

1046-5928/91 $3.00

MADE IN THE UNITED STATES OF AMERICA

This journal is printed on acid-free paper.

♾

PROTEIN EXPRESSION AND PURIFICATION

(ISSN 1046-5928)

Published bimonthly by Academic Press, Inc., 6277 Sea Harbor Drive, Orlando, FL 32887-4900

1991: Volume 2. Price: $140.00 U.S.A. and Canada; $164.00 all other countries
1992: Volume 3. Price: $140.00 U.S.A. and Canada; $164.00 all other countries
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publishers. All correspondence and subscription orders should be addressed to the office of the Publishers at 6277 Sea Harbor Drive, Orlando, FL 32887-4900. Send notices of change of address to the office of the Publishers at least 6 to 8 weeks in advance. Please include both old and new addresses. POSTMASTER: Send changes of address to *Protein Expression and Purification*, 6277 Sea Harbor Drive, Orlando, FL 32887-4900.

IDW AF / IDW / 4

CC

Sterne Kessler Goldstein Fox
ATTORNEYS AT LAW

April 21, 2005

WRITER'S DIRECT NUMBER:
(202) 772-8560
INTERNET ADDRESS:
RESMOND@SKGF.COM

Mail Stop Appeal Brief - Patents

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Re: U.S. Utility Patent Application
Application No. 09/064,057; Filed: April 22, 1998
For: **Recombinant Methods for Making Reverse Transcriptases and Mutants Thereof**
Inventors: Gerard *et al.*
Our Ref: 0942.4330002/RWE/HCC

Sir:

Transmitted herewith for appropriate action are the following documents:

1. Form PTO-2038 Credit Card Payment Form for $500.00;
2. Brief on Appeal Under 37 C.F.R. § 41.37 along with Exhibits 1-4; and
3. One (1) return postcard.

It is respectfully requested that the attached postcard be stamped with the date of filing of these documents, and that it be returned to our courier. In the event that extensions of time are necessary to prevent abandonment of this patent application, then such extensions of time are hereby petitioned.

The U.S. Patent and Trademark Office is hereby authorized to charge any fee deficiency, or credit any overpayment, to our Deposit Account No. 19-0036.

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

Robert W. Esmond

Robert W. Esmond
Attorney for Applicants
Registration No. 32,893

RWE/HCC/AWL:cpn
Enclosures

389322_1.DOC